Offshoring and the Role of Trade Agreements*

Pol Antràs
Harvard University and NBER

Robert W. Staiger
Stanford University and NBER

October 25, 2009

Abstract

The rise of offshoring of intermediate inputs raises important questions for commercial policy. Do the distinguishing features of offshoring introduce novel reasons for trade policy intervention? Does offshoring create new problems of global policy cooperation whose solutions require international agreements with novel features? In this paper we provide answers to these questions, and thereby initiate the study of trade agreements in the presence of offshoring. Our findings indicate that the rise of offshoring is likely to complicate the task of trade agreements for two reasons: first, because the mechanism by which countries can shift the costs of intervention on to their trading partners is more complicated in the presence of offshoring and extends to a wider set of policies than is the case when offshoring is not present, implying that the agreements themselves must extend to a wider set of policies as well; and second, because the underlying problem that a trade agreement must address in the presence of offshoring varies with the political preferences of member governments. As a consequence, the increasing prevalence of offshoring is likely to make it increasingly difficult for governments to rely on simple concepts and general rules – such as market access, reciprocity and non-discrimination – to help them solve their trade-related problems.

*We thank Costas Arkolakis, Kyle Bagwell, Meredith Crowley, Elhanan Helpman, Nuno Limao, Giovanni Maggi, Alan Sykes and seminar participants at Northwestern University, Stanford University, the NBER ITI Winter Meeting and the WTO and International Trade Research Conference at Brandeis University for helpful comments and discussions. Staiger gratefully acknowledges financial support from the NSF (SES-0518802). Eduardo Morales provided superb research assistance.

1 Introduction

International trade in intermediate inputs is a prominent feature of the world economy. Using OECD input-output tables, Ramanarayanan (2006) concludes that in the late 1990s imports of intermediate goods comprised between forty and sixty percent of total merchandise imports for a large number of OECD countries. Similarly, a thorough examination of highly disaggregated trade data led Yeats (2001) to conclude that intermediate input trade accounted for roughly thirty percent of world trade in manufacturing goods in 1995. Furthermore, several authors have noted that the share of intermediate inputs in world trade appears to have increased significantly in recent years.[1] This surge in intermediate input trade seems to have been accompanied by a parallel increase in the share of differentiated products in the total volume of world trade, and an associated fall in the share of goods traded on organized exchanges (also referred to as " homogeneous" goods). For instance, Rauch (1999), reports that the share of goods traded on organized exchanges in the total volume of trade fell from 27.2% in 1980 to 12.6% in 1990.[2]

Recent developments in international trade theory have attempted to bridge the apparent gap between the characteristics of international trade in the data and the standard representation of these trade flows in terms of homogeneous final goods in neoclassical trade theory. One branch of this new literature has focused on incorporating input trade in otherwise standard models with homogeneous goods, perfectly competitive markets and frictionless contracting.[3] Another branch of the literature has emphasized that modeling "offshoring" as simply an increase in the fragmentation of production across countries misses important characteristics of intermediate input trade.[4] Prominent among these characteristics is that intermediate input purchases tend to be associated with significant lock-in effects for both buyers and sellers. For example, intermediate inputs are frequently customized to the needs of their intended buyers and hence embody a disproportionate amount of relationship-specific investments, which may be hard to recoup when transacting with alternative parties. Moreover, offshoring often involves the costly search for suitable foreign suppliers or foreign buyers, which makes separations costly and thereby provides another source of lock in. Because contracts involving international transactions are especially hard to enforce, the cross-border exchange of intermediate inputs cannot generally be governed by the same contractual safeguards that typically accompany similar exchanges occurring within borders. As a consequence, these lock-in effects naturally result in prices that are determined by bilateral negotiations between suppliers and buyers, and that are therefore not (fully) disciplined by market-clearing considerations. We view the recent decline in the importance in world trade of goods traded on organized

[1] See, for instance, Feenstra and Hanson (1996b), Feenstra (1998), Campa and Goldberg (1997), Hummels, Ishii and Yi (2001) and Borga and Zeile (2004).

[2] For the case of China, the share of exports of homogeneous goods in total exports fell from 58% in 1984 to 11% in 2000 (see Kang, 2008).

[3] See for instance Feenstra and Hanson (1996a), Jones (2000), Deardorff (2001), Antràs, Garicano and Rossi-Hansberg (2006), or Grossman and Rossi-Hansberg (2008). Antràs and Rossi-Hansberg (2008) review this literature.

[4] Theoretical developments include the work of McLaren (2000), Grossman and Helpman (2002, 2005), Antràs (2003, 2005), and Antràs and Helpman (2004). See Helpman (2006) and Antràs and Rossi-Hansberg (2008) for surveys of this literature.

exchanges as a manifestation of these distinctive features of offshoring (though this decline can be attributed to other factors as well).[5]

The rise of offshoring raises important questions for commercial policy. Do the distinguishing features of offshoring introduce novel reasons for trade policy intervention? Does offshoring create new problems of global policy cooperation whose solutions require international agreements with novel features? Can trade agreements that are designed to address problems that arise when trade predominantly takes the form of the exchange of final goods be expected to perform in a world where offshoring is prevalent?

In this paper we provide answers to these questions, and thereby initiate the study of trade agreements in the presence of offshoring. We adopt the simplest setting that can capture the main features of offshoring that we wish to study, and then later show that our main points are robust to a variety of generalizations. We consider two "small" countries, Home and Foreign, who face a fixed world price for a single homogeneous final good. Production of the final good requires a customized input; all final-good producers are located in Home; and all input suppliers are located in Foreign.[6] Contracts between suppliers and producers are incomplete, and so the terms of exchange between input suppliers and final-good producers are determined by bargaining ex post (after investment in input supply has already been determined). Finally, we abstract initially from political economy concerns, and take real aggregate income as our measure of national and world welfare.

From this starting point, we investigate the role of trade policies. We assume that each country can apply taxes/subsidies to trade in the input and/or the final good. We first consider the case for free trade in this environment. As might be expected, the distortions associated with international hold up create an activist role for policy intervention from the perspective of world welfare. Intuitively, the combination of the lock-in effect (created by relationship-specific investments) and the incompleteness of contracts results in an international hold-up problem that leads to an inefficiently low volume of input trade across countries under free trade. It is therefore natural that trade policies which encourage input trade volume can serve as a substitute for more standard contractual safeguards available in domestic transactions and can thereby help bring countries closer to the efficiency frontier. Importantly, though, the mechanism by which trade policies affect input trade volumes in this environment is by altering the conditions of ex-post *bargaining* between foreign suppliers and domestic producers, and is therefore distinct from the standard manner that trade volumes respond to trade policy intervention in the commercial policy literature (through changes in domestic demand and/or competing domestic supply and their implications for international market-clearing conditions). Nevertheless, we show that an appropriate choice of input trade subsidies, combined with free trade in final goods, can fully resolve the international hold-up problem and allow countries to attain the first-best.

[5] The recent empirical studies of Feenstra and Hanson (2005), Yeaple (2006), Levchenko (2007), Nunn (2007), and Nunn and Trefler (2008) also substantiate the empirical relevance of the aforementioned non-standard features of offshoring.

[6] For the most part, we illustrate the emergence of a lock-in effect by appealing to customization of inputs. We will demonstrate, however, that our model can be interpreted as a reduced form of a dynamic model where this lock-in effect stems from search frictions even when inputs are not specialized.

We next ask whether the Nash equilibrium policy choices of governments coincide with the internationally efficient policies. We find that they do not, and we identify two dimensions of international inefficiency that arise under Nash policies: an inefficiently low input trade volume; and an inefficiently low price of the final good in the Home market. Intuitively, trade policy serves a dual role in this environment. On the one hand, as indicated above, subsidies to the exchange of intermediate inputs can serve as a substitute for more standard contractual safeguards available in domestic transactions and can thus increase the volume of input trade toward its efficient level. On the other hand, input trade taxes can be used to redistribute surplus across countries, thereby shifting some of the cost of intervention on to trading partners. For instance, although an export tax may reduce the incentive of Foreign suppliers to invest, in their ex-post bargaining these suppliers will be able to pass part of the cost of the tax on to final-good producers in Home. Moreover, we show that the home government will also distort trade in the final good away from its free-trade level in order to reduce the domestic final-good price and further shift bargaining surplus from Foreign input suppliers to Home final-good producers in this fashion. There is hence a basic tension that each government faces in its unilateral trade policy choices between correcting the hold-up problem and capturing surplus from/shifting costs onto its trading partner, and this tension prevents governments from making internationally efficient policy choices in the Nash equilibrium.

We then turn to the role of trade agreements in this setting. Our description of the Nash inefficiencies above raises the natural question whether these inefficiencies can be attributed to terms-of-trade manipulation, as in the "terms-of-trade theory" that characterizes much of the existing trade agreements literature (see Bagwell and Staiger, 2002, Ch. 2 for a review). Providing an answer to this question is important not only for the purpose of positioning our results within the existing literature, but also for the purpose of understanding the design features of a trade agreement that could help governments achieve the efficiency frontier when offshoring is a prevalent feature of the world economy. In answer to this question, we find that, provided governments seek to maximize national income with their trade policy choices, a terms-of-trade interpretation can indeed be given to the inefficiencies associated with Nash policy choices in the Benchmark Model, and we further show that the GATT/WTO principle of "reciprocity," which has been shown to be an attractive feature of a trade agreement in a wide variety of settings (see, for example, Bagwell and Staiger, 1999, 2001a, and most recently 2009a,b), continues to exhibit attractive features in the presence of offshoring. But as we demonstrate, there is one central feature of the GATT/WTO system whose desirable properties do not carry over to a setting such as the Benchmark Model in which offshoring is present: this feature is the GATT/WTO focus on "market access," according to which negotiations are narrowly focused on the tariffs that are applied directly to the trade flows in question, and where the concept of "nullification or impairment" is then utilized to protect the implied access from subsequent unilateral adjustments that a government might wish to make to its wider portfolio of policies. This leads to our first broad conclusion: the rise in offshoring is likely to complicate the task of trade agreements, because in the presence of offshoring the mechanism by which countries can shift the costs of intervention on to their trading partners is more complicated and extends to a *wider set of policies* than is the case when offshoring is not present, creating the need for negotiations that then must extend to a wider set of policies as well.

We next introduce the possibility that governments are motivated in part by political economy/redistributive concerns. We show that the introduction of political economy motives into our model can eliminate unrealistic policy predictions (e.g., convert import subsidies to import taxes and export taxes to export subsidies), but we confirm that the implications of offshoring for the comparison between Nash and efficient trade policies as described above is preserved. More specifically, we establish that sufficiently politically motivated governments will adopt import tariffs and export subsidies in the Nash equilibrium, but we show that Nash policies still imply inefficiently low input trade volume and an inefficiently low price of the final good in the Home market.

But when we ask whether these inefficiencies can still be attributed to terms-of-trade manipulation once the political motivations of governments are introduced, we find a surprising answer: a new inefficiency arises in the Nash equilibrium that cannot be given a terms-of-trade interpretation, and which instead relates to the budgetary consequences of export promotion. This finding is in stark contrast to the predictions of the terms-of-trade theory, where the presence or absence of political economy motives has no impact on the underlying problem that a trade agreement must solve. We thus conclude that if governments have political economy motivations, an effective trade agreement must serve two roles in the presence of offshoring: it must provide governments with an avenue of escape from a terms-of-trade driven Prisoners' Dilemma; and it must coordinate the setting of policies across countries so as to reduce the budgetary burden of export promotion programs for purposes of redistribution. This leads to our second broad conclusion: the rise in offshoring is likely to complicate the task of trade agreements, because in the presence of offshoring the underlying problem that a trade agreement must address *varies* with the political preferences of member governments. As a consequence, under the view that governments operate in the presence of important political economy forces, the increasing prevalence of offshoring is likely to make it increasingly difficult for governments to rely on simple and general rules – such as reciprocity and non-discrimination – to help them solve their trade-related problems.

Our paper is related to several literatures. First, as emphasized above, by exploring the role of trade agreements in a model with intermediate input trade and in an environment with relationship-specific investments and incomplete contracting, we complement and extend an established literature on international trade agreements (see Bagwell and Staiger, 2009a,b for recent contributions to this literature that explore related themes). In suggesting a novel rationale for trade agreements, our paper also complements the recent papers of Ossa (2008) and Mrazova (2009). Second, by considering endogenous trade policy choices in this environment, we complement and extend a recent literature that has begun to study the impacts of (exogenous) tariffs on international hold-up problems. Ornelas and Turner (2008a) develop a model in which import tariffs on intermediate inputs are shown to aggravate the hold-up problem in international vertical relationships, with the implication that trade liberalization may lead to a larger increase in trade flows than in standard models. Ornelas and Turner do not however study optimal trade policies or the possibility of trade agreements in their framework.[7,8] McLaren (1997) studies the desirability of announcing a future trade liberalization in a model where producers incur sunk costs to service foreign markets, but his framework emphasizes commitment problems from which we completely abstract.[9]

Finally, there is a large literature proposing a variety of mechanism-design resolutions to the hold-up inefficiencies caused by incomplete contracts. These resolutions however generally rely on the ability of parties to commit not to renegotiate an initial contract and also on the existence of a third party that can enforce off-the-equilibrium-path penalties.[10] We view our international context as one in which these alternative resolutions of the hold-up problem are naturally more problematic, and thus trade taxes and subsidies may be particularly useful in resolving these inefficiencies. For this same reason, we find it natural to simplify our model in a way that avoids completely any source of domestic hold-up inefficiencies.[11]

The rest of the paper is organized as follows. In section 2, we develop a Benchmark Model that introduces the international hold-up problem and illustrates the role of active second-best trade policies. In section 3, we consider Nash equilibrium policy choices when governments maximize national income and show that Nash policies are inefficient. Section 4 explores the role and design of trade agreements in this setting, while section 5 extends the analysis of the Benchmark Model to include political economy motives. In section 6, we consider a variety of further extensions of the model. We offer some concluding remarks in section 6.

2 The Benchmark Model

We begin this section by describing a benchmark two-small-country trade model in which final-good producers in the home country import inputs from suppliers in the foreign country. We refer to this model as the *Benchmark Model*. While simple and special along a number of dimensions, the Benchmark Model is meant to highlight the essential features of the basic international hold-up problem which arises under free trade. After presenting the setup and characterizing the free-trade equilibrium, we derive the (second-best) trade policies that maximize world welfare.

2.1 Setup

We consider a world of two small countries, Home (H) and Foreign (F), and a large rest-of-world whose only role in the model is to fix the price at which a final good 1 is available to H and F on world markets (the direction of trade in good 1 is not specified and is immaterial). Consumer

[7] The independent paper of Ornelas and Turner (2008b) does begin to explore the welfare implications of tariffs in this kind of environment, but the they do not consider the role of trade agreements.

[8] Similarly, Antràs and Helpman (2004) and Diez (2008) study the effect of trade frictions on the choice of organizational form of firms contemplating offshoring, but they also treat trade frictions as exogenous.

[9] Yarbrough and Yarbrough (1992) also emphasize commitment problems associated with trade relationships that involve substantial relationship- (or market-) specific investments, but they focus on how these issues affect the choice between unilateral liberalization, bilateral agreements and multilateral agreements.

[10] Bolton and Dewatripont (2005, Chapter 12) review the insights and limitations of this literature.

[11] In related work, Rosenkranz and Schmitz (2007) show that, in a closed economy setup, a government can use taxation to alleviate the hold-up problem between domestic buyers and sellers.

preferences are identical in H and F and given by

$$U^j = c_0^j + u\left(c_1^j\right), \tag{1}$$

where c_i^j is consumption of good $i \in \{0,1\}$ in country $j \in \{H,F\}$, and where $u' > 0$ and $u'' < 0$. Good 0, which we take to be the numeraire, is assumed to be costlessly traded and available in sufficient quantities that it is always consumed in positive amounts in both H and F. Good 1 is produced with a customized intermediate input x according to the production function $y(x)$, with $y(0) = 0$, $y'(x) > 0$ and $y''(x) < 0$. By choice of units for measuring the quantity of good 1, we set its (fixed) price on world markets equal to 1. For now we assume that trade in good 1 is free, so that its price is equal to 1 everywhere in the world.

Notice that the concavity of $y(x)$ implies $y(x)/x > y'(x)$ for $x > 0$. We impose as well an additional condition on the curvature properties of $y(x)$:

$$\frac{y(x)}{x} - y'(x) + xy''(x) < 0 \text{ for } x > 0. \tag{2}$$

As we establish later, this condition ensures that the home country improves its terms of trade when it imposes a tariff on imports of the intermediate input x, and it thereby rules out the Lerner Paradox in the Benchmark Model.[12]

We suppose that the home country H is inhabited by a unit measure of producers of the final good 1, while the foreign country F is inhabited by a unit measure of suppliers of the intermediate input x. Hence, to produce the final good 1, producers in H must import inputs from suppliers in F. Suppliers in F tailor their inputs specifically to the needs of a final-good producer in H and, for simplicity, these inputs are assumed to be useless to alternative final-good producers. We assume that the marginal cost of input production in F is constant and, through choice of the units in which inputs are measured, we normalize it to 1. For now, we also assume that trade in x is free.

We next turn to focus on the nature of the bilateral relationship between a final-good producer in H and an input supplier in F, which comprises the essence of the model. We adopt a setting of incomplete contracts between final-good producers and input suppliers. In our Benchmark Model, contractual incompleteness can be rationalized in the following simple way. Following Grossman and Helpman (2002) and Antràs (2003), we assume that, when investing in the supply of x, the supplier can choose between manufacturing a high-quality or a low-quality input, and the latter can be produced at lower cost but is useless to final-good producers. The quantity of x is observable to everyone and therefore verifiable by third-parties, but we assume that the quality of x is only observable to the supplier and producer in the particular bilateral relationship, and so quality-contingent contracts are not available. Although parties could still sign a contract specifying a price and a quantity, if they did so, the supplier would always have an incentive to produce the low quality input (at lower cost) and still receive the same contractually stipulated price.

Hence, in this environment, no (enforceable) contracts are signed between suppliers and pro-

[12] In order to ensure that the second-order conditions are met, we will later impose additional assumptions on $y(x)$.

ducers prior to the initial supplier investment decisions. And without an initial contract, the price at which each supplier in F sells its inputs to a producer in H is then decided ex-post (through bargaining) once quality has been chosen. We follow the bulk of the literature in assuming that the bargained price is determined through symmetric Nash bargaining. Because parties have symmetric information at the bargaining stage, ex-post efficiency ensures that low-quality production will never be chosen by an input supplier in equilibrium, and so only high-quality inputs are produced: as a result, the input-quality dimension of the model can be kept in the background henceforth.

We now describe the structure of the bilateral producer-supplier relationship in detail. We assume that all agents have an ex-ante zero outside option. The sequence of events is as follows:

stage 1. The unit measure of producers in H and suppliers in F are randomly matched, producing a unit measure of matches. Each agent decides whether to stay with his match or exit the market. In the former case, the producer provides the supplier with a list of customized input specifications. In the latter case, each agent obtains his ex-ante outside option (equal to zero).

stage 2. Each supplier decides on the amount x of customized input to be produced (at marginal cost of 1).

stage 3. Each producer-supplier pair bargains over the price of the intermediate input (we assume symmetric Nash bargaining).

stage 4. Each producer in H imports x from its partner-supplier and produces the final good with the acquired x, and payments agreed in *stage* 3 are settled.

This 4-stage game generates the simple hold-up problem that forms the heart of our analysis. A number of features of this setup are worth noting at this point.

First, we rule out the use of ex-ante (*stage*-1) lump-sum transfers between producers and suppliers. The possibility of these transfers is particularly hard to defend in the international context that we study, where such transfers and the obligations associated with them might be difficult to enforce. In section 6, however, we show that our main results are robust to allowing for these transfers. Second, we assume a frictionless matching process in *stage* 1 to keep our Benchmark Model simple: in section 6 we introduce (ex-ante) search frictions. Third, we assume symmetric Nash bargaining in *stage* 3. This helps to keep the number of parameters to a minimum in our Benchmark Model and allows us to focus on the main points. However, we relax this assumption in section 6, where we consider general bargaining power. Fourth, the role of the specificity of input x is to pin down the outside options of the producer and the supplier should their *stage*-3 bargaining break down. In our Benchmark Model we take an extreme view of the degree of specificity, so that the breakup of a bargaining pair in *stage* 3 would result in a zero outside option for both producer and supplier. We also relax this assumption in section 6, where we introduce a secondary market for inputs. As argued in the Introduction, we could altogether dispense with the assumption of specificity of inputs by introducing (ex-post) search frictions, which again would drive a wedge between the value of remaining in a match and the value of dissolving that match. In fact, our

Benchmark Model is isomorphic to a model with extreme (ex-post) search frictions, in which a separation implies that each party finds an alternative trading partner with probability 0. Our less extreme framework in section 6 is isomorphic to a model with less extreme search frictions.

Finally, we note that production efficiency requires that the customized input is produced at a level x^E which satisfies

$$y'\left(x^E\right) = 1, \tag{3}$$

and thereby equates the marginal revenue generated from an additional unit of the input (recall that the price of the final good is fixed by world markets and equal to 1 under free trade) with the marginal cost of producing an additional unit of the input (which is constant and normalized to 1).

2.2 Free Trade Equilibrium

We now characterize the subgame perfect equilibrium of the 4-stage game described above. The characterization follows very simply from a few key observations. We consider a representative producer in H and supplier in F that are matched in *stage* 1.

First, if the producer uses the supplier's input to produce the final good in *stage* 4, its revenue is given by $y(x)$. Second, as observed in the previous section, the outside options of both the producer and the supplier in their *stage*-3 Nash bargain are 0, and hence the quasi-rents over which the producer and supplier bargain in *stage* 3 (recall that the cost of producing x is sunk at this point) are $y(x)$. Therefore, in the symmetric Nash bargain of *stage* 3, the final-good producer in H and the input supplier in F both obtain a payoff of $\frac{1}{2}y(x)$.

Next, rolling back to *stage* 2, observe that the input supplier chooses x to maximize $\frac{1}{2}y(x) - x$, so the optimal quantity $\hat{x}$ of input satisfies

$$y'(\hat{x}) = 2. \tag{4}$$

Given the concavity of $y(x)$, it is clear from a comparison of (4) with (3) that $\hat{x} < x^E$. This is the under-investment associated with the hold-up problem, and it reflects the fact that the producer and supplier bargain over the price of the input *after* the supplier has already sunk investment in input supply.

Finally, consider *stage* 1. If the producer hands the supplier a list of customized input specifications, the producer anticipates obtaining a payoff equal to

$$\pi^H = \frac{1}{2}y(\hat{x}),$$

which exceeds the payoff he would obtain by not providing the specifications (recall that the ex-ante outside option of producers is equal to 0). Similarly, by agreeing to form a partnership with the home producer, the supplier anticipates obtaining a payoff of

$$\pi^F = \frac{1}{2}y(\hat{x}) - \hat{x},$$

which also exceeds his ex-ante outside option.[13] In sum, no separations will occur at *stage* 1. Note also that the sum of payoffs of the two parties is equal to $y(\hat{x}) - \hat{x}$, which is strictly less than the sum of payoffs that would obtain when investment is chosen at the efficient level x^E defined by (3).

Now consider the measure of social welfare in each country implied by our Benchmark Model. With our assumption of quasilinear preferences, this measure is given by consumer surplus plus profits plus trade tax revenue (the latter being zero under free trade).[14] Using (1), we have that country j's demand for good 1 is given by $D_1(p_1^j) \equiv u'^{-1}(p_1^j)$, with consumer surplus then defined as $CS^j(p_1^j) \equiv \int_{p_1^j}^{\bar{p}} D_1(p)dp$ where $\bar{p}$ is the "choke" price for country j's demand of good 1. World aggregate welfare may then be represented by

$$W^W = W^H + W^F = CS^H(1) + CS^F(1) + \pi^H + \pi^F = CS^H(1) + CS^F(1) + y(\hat{x}) - \hat{x},$$

which is strictly lower than world welfare in the presence of production efficiency because $y(\hat{x}) - \hat{x} < y\left(x^E\right) - x^E$. We summarize this discussion with:

Proposition 1 *In the Benchmark Model, a hold-up problem exists under free trade, leading to an inefficiently low volume of input trade ($\hat{x} < x^E$).*

Proposition 1 records the existence of a basic international hold-up problem that arises in the presence of free trade. At this point, there are a variety of mechanism-design resolutions to the hold-up inefficiencies caused by incomplete contracts that we might consider. However, we view our international context as one in which these mechanism-design resolutions are naturally more problematic because they generally rely on the ability of parties to commit not to renegotiate an initial contract and also on the existence of a third party that can enforce off-the-equilibrium-path penalties. In this light, trade taxes and subsidies may be particularly useful as an alternative route to resolving these inefficiencies. We therefore next turn to consider trade intervention as a possible means of alleviating the hold-up problem.

2.3 Second-Best Trade Policy

In this section, we explore the possible beneficial role of trade policy in this distorted economy. To this end, we let τ_x^H denote the trade tax imposed by H on imports of the input x (positive if an import tariff, negative if an import subsidy) defined in specific terms, and we let τ_x^F be the analogous trade tax imposed by F (positive if an export tax, negative if an export subsidy). Furthermore, we let τ_1^H denote the trade tax imposed by H on the home country's trade in the final good 1 (positive if an import tariff or export subsidy, negative if an import subsidy or export tax) also defined in specific terms. Observe that the price of the final good 1 in H is now given

[13] Given the concavity of $y(x)$, we have $\frac{1}{2}y(\hat{x}) - \hat{x} \geq \frac{1}{2}\hat{x}y'(\hat{x}) - \hat{x} = 0$.

[14] Strictly speaking, social welfare should also include a term related to income earned by other factors of production (say labor) in the economy. Nevertheless, it is straightforward to close the model in a way that makes this term independent of policies in sector 1 (see, for instance, Grossman and Helpman, 1994). Henceforth, we simply ignore this term.

by $p_1^H = 1 + \tau_1^H$, whereas the price of the input x continues to be determined by Nash bargaining between producers and suppliers (though trade taxes may affect this negotiated price).[15]

How does the introduction of these trade taxes affect the equilibrium characterized in the previous section? To explore this question, we first consider the case of second-best trade policies, that is, the set of policies that maximize aggregate world welfare (subject to the contractual frictions in producer-supplier relationships). More specifically, we introduce the following *stage* 0 which occurs prior to *stage* 1 of the 4-stage game described in section 2.1:

stage 0. A social planner selects a home-country trade tax τ_1^H on the final good 1, a home-country import tax τ_x^H on home imports of the input x, and a foreign-country export tax τ_x^F on foreign exports of the input x.

After the social planner has selected these import tariffs/subsidies in *stage* 0, the sequence of events is as outlined in section 2.1 (with trade taxes collected at the time of importation and production/sales in *stage* 4).

Consider now how these trade policy choices in *stage* 0 affect the equilibrium outcome of the game. In their *stage*-3 bargaining, if the producer and supplier reach an agreement they stand to obtain a joint payoff of (recalling again that the cost of producing x is sunk at that point)

$$\left(1+\tau_1^H\right) y\left(x\right) - \left(\tau_x^H + \tau_x^F\right) x.$$

A positive import tariff or export subsidy on the final good ($\tau_1^H > 0$) raises the joint surplus of the producer and supplier because it raises the price at which the final good is sold in H. Conversely, a positive import tariff ($\tau_x^H > 0$) or export tax ($\tau_x^F > 0$) on inputs reduces the joint surplus of the producer and supplier because it transfers part of the surplus to governments.

If the producer and the supplier do not reach an agreement, each is again left with a zero outside option. Hence, both parties obtain a payoff equal to $\frac{1}{2}\left(\left(1+\tau_1^H\right) y\left(x\right) - \left(\tau_x^H + \tau_x^F\right) x\right)$ in the Nash bargain of *stage* 3, and the input supplier's choice of x in *stage* 2 must then satisfy[16]

$$\frac{1}{2}\left(1+\tau_1^H\right) y'\left(\hat{x}\right) = 1 + \frac{1}{2}\left(\tau_x^H + \tau_x^F\right). \tag{5}$$

It is clear from (5) that $\hat{x}$ is increasing in τ_1^H and decreasing in τ_x^H and τ_x^F. Intuitively, incomplete contracting leads to rent-sharing between the producer and supplier, and the latter's incentives to invest tend to be higher whenever the surplus from investment is higher, that is when τ_1^H is higher and when τ_x^H or τ_x^F are lower. We will see in later sections that the positive dependence of $\hat{x}$ on τ_1^H

[15] We could also allow for a final-good trade tax τ_1^F in the foreign country, but it is intuitively clear (and is easily shown) that there will be no incentive to use such an instrument, since such trade taxes could only alter the local price of good 1 in F (owing to F's small size on world markets) and that price has no impact on the hold-up problem between F's input suppliers and H's final good producers.

[16] Implicit in our discussion is the assumption that $\frac{1}{2}\left(\left(1+\tau_1^H\right) y\left(x\right) - \left(\tau_x^H + \tau_x^F\right) x\right) > 0$, so that the Nash bargain payoff beats each party's outside option. It is straightforward to show that this is true in our Benchmark Model for the relevant values of home and foreign policies.

and negative dependence of $\hat{x}$ on τ_x^H and τ_x^F hold for a variety of specifications of the game played between the producer and supplier.

At *stage* 1, the final-good producer in H anticipates a payoff equal to

$$\pi^H = \frac{1}{2}\left(\left(1+\tau_1^H\right) y\left(\hat{x}\right) - \left(\tau_x^H + \tau_x^F\right)\hat{x}\right), \tag{6}$$

while the supplier in F expects a payoff equal to

$$\pi^F = \frac{1}{2}\left(\left(1+\tau_1^H\right) y\left(\hat{x}\right) - \left(\tau_x^H + \tau_x^F\right)\hat{x}\right) - \hat{x}, \tag{7}$$

where $\hat{x}$ is implicitly defined by (5). As a result, welfare in H, inclusive of tax revenue, is given by

$$W^H = CS^H(1+\tau_1^H) + \pi^H + \tau_1^H\left[D_1(1+\tau_1^H) - y\left(\hat{x}\right)\right] + \tau_x^H\hat{x},$$

while welfare in F is

$$W^F = CS^F(1) + \pi^F + \tau_x^F\hat{x}.$$

We now seek to characterize the set of trade policy choices that maximize world welfare. Formally, we are seeking the triplet $\left(\tau_1^{HE}, \tau_x^{HE}, \tau_x^{FE}\right)$ that maximizes (see (6) and (7)):

$$W^W = W^H + W^F = CS^H(1+\tau_1^H) + CS^F(1) + \left(1+\tau_1^H\right) y\left(\hat{x}\right) - \hat{x} + \tau_1^H\left[D_1(1+\tau_1^H) - y\left(\hat{x}\right)\right],$$

subject to $\hat{x}$ being given by (5).[17] Notice that for a given value of the sum $\tau_x^H + \tau_x^F$, the individual values of τ_x^H and τ_x^F have no effect on world welfare. This implies that the second-best policies will only pin down an aggregate input trade tax $\tau_x \equiv \tau_x^H + \tau_x^F$. The efficient policies τ_1^{HE} and τ_x^E are then determined by the following first-order conditions of the problem above:[18]

$$\begin{aligned} \frac{\partial W^W}{\partial \tau_1^H} &= \tau_1^H \frac{\partial D_1}{\partial p_1^H} + \left[y'\left(\hat{x}\right) - 1\right]\frac{\partial \hat{x}}{\partial \tau_1^H} = 0, \text{ and} \\ \frac{\partial W^W}{\partial \tau_x} &= \left[y'\left(\hat{x}\right) - 1\right]\frac{\partial \hat{x}}{\partial \tau_x} = 0. \end{aligned} \tag{8}$$

The first-order conditions in (8) are instructive. Recalling from (5) that $\partial\hat{x}/\partial\tau_1^H > 0$, it is clear from (8) that the optimal choice of τ_1^H is strictly positive, provided that $\left[y'\left(\hat{x}\right) - 1\right] > 0$ which by (3) implies that $\hat{x} < x^E$: this suggests that an import tariff or export subsidy on trade in the final good 1 could raise welfare in the world, by increasing $\hat{x}$ toward x^E and thereby helping to ameliorate the hold-up problem at the cost of lost consumer surplus. However, recalling from (5) that $\partial\hat{x}/\partial\tau_x < 0$, it is clear from (8) that the optimal choice of τ_x must ensure that $\left[y'\left(\hat{x}\right) - 1\right] = 0$, thereby achieving productive efficiency: there is no associated loss in consumer surplus when the tariff on imported inputs τ_x is used to increase $\hat{x}$, and the optimal choice of τ_x therefore solves

[17] It is the presence of this constraint that leads us to refer to $\left(\tau_1^{HE}, \tau_x^{HE}, \tau_x^{FE}\right)$ as *second-best* trade policy choices, although we shall show that these policy choices lead to an attainment of the first-best welfare level.

[18] It is easily checked that second-order conditions are satisfied (see Appendix A.1).

completely the hold-up problem and achieves productive efficiency. This in turn leaves no reason for government intervention with regard to trade in the final good 1. Hence, the optimal choice of τ_1^H is $\tau_1^{HE} = 0$. On the other hand, the second-best policies do call for intervention with regards to input trade. In particular, from equation (5) it follows that the optimal trade tax is an input subsidy in an amount equal to $\tau_x^E \equiv \tau_x^{HE} + \tau_x^{FE} = -1$. We may thus state:

Proposition 2 *In the Benchmark Model, the second-best trade policy choices maintain free trade in the final good and subsidize importation of the input so as to solve the hold-up problem and achieve an efficient volume of input trade ($\hat{x} = x^E$).*

The intuition for Proposition 2 is simple. The hold-up problem between producers in H and suppliers in F results in a level of imported inputs which is inefficiently low. The market failure is an international one in nature, and thus it is natural that trade taxes or subsidies can serve a useful role in alleviating the inefficiency. Furthermore, although trade intervention in the final good could be used to raise the home-country price of the final good and increase the volume of imported inputs (through rent-sharing), this would come at a cost of reduced home-country consumer surplus. A subsidy to imported inputs does not reduce consumer surplus, but it nevertheless succeeds in increasing the volume of imported inputs by increasing the surplus over which the parties negotiate in the ex-post (*stage*-3) bargain. As a consequence, a subsidy to imported inputs targets just the distorted margin, and in analogy with the targeting principle (Bhagwati and Ramaswami, 1963, Johnson, 1965) is hence the optimal method of addressing the problem.

We have thus identified a novel role for trade policy intervention, namely, as a means of addressing the international hold-up problem that arises when international trade requires relationship-specific investments between domestic producers and their foreign suppliers. A natural question is whether the *unilateral* trade policy choices of both the home and foreign governments will lead to overall trade interventions that concord with the efficiency conditions outlined in Proposition 2. We tackle this issue in the next section.

3 Nash Trade Policy

In this section we characterize the Nash policies of the home and foreign governments and evaluate the potential role of trade agreements in our Benchmark Model. In order to build intuition, we first consider the unilaterally optimal trade policy choices of the home government when the foreign government follows a policy of free trade, and only later consider the possibility of foreign trade policy intervention.

3.1 Unilateral Home Policy

To characterize the unilaterally optimal trade policy choices of the home government when the foreign government follows a policy of free trade, we derive the subgame perfect equilibrium of the

Benchmark Model for the case in which *stage* 0 is as follows:[19]

stage 0. The home government H selects a trade tax τ_1^H on the final good 1, and a trade tax τ_x^H on the imported input x; the foreign government F remains passive, i.e., $\tau_x^F \equiv 0$.

Following the same steps as in the last section, and with τ_x^F set to zero at *stage* 0, we have that the final-good producer in H now obtains a *stage*-2 payoff equal to:

$$\pi^H = \frac{1}{2}\left(\left(1+\tau_1^H\right) y\left(\hat{x}\right) - \tau_x^H \hat{x}\right), \tag{9}$$

where $\hat{x}$ is now implicitly defined by

$$\frac{1}{2}\left(1+\tau_1^H\right) y'\left(\hat{x}\right) = 1 + \frac{1}{2}\tau_x^H. \tag{10}$$

With these expressions in hand, home welfare can be written as the sum of home consumer surplus, profits and tax revenue, or

$$W^H = CS(1+\tau_1^H) + \frac{1}{2}\left(\left(1+\tau_1^H\right) y\left(\hat{x}\right) - \tau_x^H \hat{x}\right) + \tau_1^H [D_1(1+\tau_1^H) - y\left(\hat{x}\right)] + \tau_x^H \hat{x}.$$

The optimal choice of τ_1^H and τ_x^H, which we denote by $\hat{\tau}_1^H$ and $\hat{\tau}_x^H$, must maximize home welfare W^H, and will hence satisfy the first-order conditions

$$\begin{aligned}
\frac{\partial W^H}{\partial \tau_1^H} &= 0 = \tau_1^H \frac{\partial D_1}{\partial p_1^H} - \frac{1}{2} y\left(\hat{x}\right) + \left[\frac{1}{2}\left(1-\tau_1^H\right) y'\left(\hat{x}\right) + \frac{1}{2}\tau_x^H\right] \frac{\partial \hat{x}}{\partial \tau_1^H}, \text{ and} \\
\frac{\partial W^H}{\partial \tau_x^H} &= 0 = \frac{1}{2}\hat{x} + \left[\frac{1}{2}\left(1-\tau_1^H\right) y'\left(\hat{x}\right) + \frac{1}{2}\tau_x^H\right] \frac{\partial \hat{x}}{\partial \tau_x^H},
\end{aligned}$$

where recall that $\hat{x}$ is given by equation (10).[20] Applying the implicit function theorem (twice) to (10) delivers $\frac{\partial \hat{x}/\partial \tau_1^H}{\partial \hat{x}/\partial \tau_x^H} = -y'\left(\hat{x}\right)$, which can be used to manipulate the above first-order conditions to obtain:

$$\begin{aligned}
\hat{\tau}_1^H &= -\frac{\frac{1}{2}\hat{x}[\frac{y(\hat{x})}{\hat{x}} - y'\left(\hat{x}\right)]}{|\partial D_1/\partial p_1^H|}, \text{ and} \\
\hat{\tau}_x^H &= -(1-\hat{\tau}_1^H) y'\left(\hat{x}\right) - \frac{\hat{x}}{\partial \hat{x}/\partial \tau_x^H}.
\end{aligned} \tag{11}$$

[19] Implied by this timing of tariff choices is the assumption that governments can make tariff commitments to the private sector. If the governments did not have this ability, then as is well known a separate commitment role for trade agreements might arise (see Bagwell and Staiger, 2002, Chapter 2, for a review of this literature). As noted in the Introduction, the particular commitment problems that governments face when trade requires relationship-specific investments are emphasized by Yarbrough and Yarbrough (1992) as providing a reason for trade agreements to exist, and by McLaren (1997) as creating the possibility of perverse negotiating outcomes. Our assumed timing permits us to abstract from the possible commitment role of trade agreements throughout this paper, so that we may focus on other issues.

[20] The second-order conditions for this problem do not reduce to simple expressions, as was the case with second-best policies. In Appendix A.1, we discuss these second-order conditions and show that they are satisfied for a simple parameterized example.

The expressions in (11) reflect an interesting logic. Part of the goal of the home government in intervening with τ_1^H and/or τ_x^H, as in the case of second-best policies, is to raise $\hat{x}$ towards its efficient level x^E. Nevertheless, the home government *does not* maximize world welfare and hence there is an offsetting leakage of surplus to the foreign supplier that must be taken into account by the home government in setting its optimal unilateral policies. This leads to two observations: first, it is not optimal for the home government to deliver the chosen $\hat{x}$ using only τ_x^H, and the setting of $\tau_1^H \neq 0$ reflects a new and independent source of international inefficiency associated with the unilateral policy choices of the home country; and second, it is not optimal for the home government to raise $\hat{x}$ all the way to its efficient level x^E.

The first observation can be understood as follows. The home government must concern itself with two tasks as it considers its policy choices. First, it must face foreign suppliers with the appropriate *marginal* incentives for investment in the supply of x so as to achieve the desired investment level $\hat{x}$. Second, the home government must also concern itself with extracting *infra-marginal* surplus from foreign suppliers through the use of trade policy instruments.

With its two tariff instruments τ_1^H and τ_x^H, the home government can extract inframarginal foreign surplus with adjustments in τ_1^H and τ_x^H that hold $\hat{x}$ fixed according to (10), so that $d\tau_x^H(\tau_1^H)/d\tau_1^H = -\frac{\partial\hat{x}/\partial\tau_1^H}{\partial\hat{x}/\partial\tau_x^H} = y'(\hat{x})$, and can extract foreign surplus in this fashion at the rate

$$\frac{dW^F(\tau_1^H, \tau_x^H(\tau_1^H), \hat{x})}{d\tau_1^H}|_{d\hat{x}=0} = \frac{1}{2}\hat{x}\left[\frac{y(\hat{x})}{\hat{x}} - y'(\hat{x})\right]. \tag{12}$$

Evidently, with the concavity of $y(x)$ implying $[\frac{y(\hat{x})}{\hat{x}} - y'(\hat{x})] > 0$, it follows from (12) that for any given level of $\hat{x}$, additional surplus can be extracted from the foreign country by reducing τ_1^H and accompanying this with a reduction in τ_x^H which preserves the level of $\hat{x}$. Intuitively, while we have seen that a positive final-good tariff τ_1^H *could* be used to induce greater investment from foreign suppliers, from the perspective of foreign surplus extraction it is an inferior method for doing so relative to a subsidy to imported inputs τ_x^H, because τ_1^H must work through the final-good production function $y(x)$ – which is concave – and this creates more infra-marginal surplus for foreign suppliers relative to τ_x^H, which works directly (and linearly) through import volume x.

What, then, prevents the home country from lowering τ_1^H and τ_x^H in this fashion indefinitely, until all of the surplus has been extracted from foreign suppliers? The impact on home-country welfare of these tariff changes is given by

$$\frac{dW^H(\tau_1^H, \tau_x^H(\tau_1^H), \hat{x})}{d\tau_1^H}|_{d\hat{x}=0} = \tau_1^H \frac{\partial D_1^H}{\partial p_1^H} - \frac{1}{2}\hat{x}\left[\frac{y(\hat{x})}{\hat{x}} - y'(\hat{x})\right]. \tag{13}$$

As (13) makes clear, what eventually stops this process of foreign surplus extraction is the growing home-country final-good demand distortions that are associated with $\tau_1^H < 0$.

It is for these reasons that (11) implies $\hat{\tau}_1^H < 0$: in words, it is unilaterally optimal for the home government to utilize trade policy to *distort downward* the price of the final good 1 in the home market (through either an import subsidy or an export tax on the final good) as a means of extracting bargaining surplus from foreign suppliers. Finally, recalling that $W^W = W^H + W^F$, note that (12) and (13) together imply

$$\frac{dW^W}{d\tau_1^H}|_{d\hat{x}=0} = \tau_1^H \frac{\partial D_1^H}{\partial p_1^H},$$

which is strictly positive for $\tau_1^H < 0$: in words, setting $\tau_1^H < 0$ is inefficient from the point of view of aggregate world welfare for *any* level of $\hat{x}$. Hence, our model identifies a new and independent source of international inefficiency when the home country sets its tariffs unilaterally: the attempt to extract bargaining surplus from foreign suppliers by distorting the home market price of the final good.

The second observation above, that it is not optimal for the home country to raise $\hat{x}$ all the way to its efficient level x^E, can be confirmed by considering the expression for $\hat{\tau}_x^H$ in (11). This expression is of indeterminate sign, indicating that $\hat{\tau}_x^H$ can now be either negative (an import subsidy on inputs of x) or positive (an import tariff on inputs of x): this reflects the tension that arises for the home-country government between correcting the hold-up problem and capturing surplus from the foreign input supplier, a tension that was absent in the choice of second-best policies in section 2.3. To show formally that the home government will not raise $\hat{x}$ to the efficient level, we substitute (10) into the expression for $\hat{\tau}_x^H$ in (11) and simplify to obtain

$$y'(\hat{x}) = 1 - \frac{1}{2} \frac{\hat{x}}{\partial \hat{x} / \partial \tau_x^H} > 1, \tag{14}$$

which implies that $\hat{x} < x^E$. Hence, at least when the foreign government remains passive, it is unilaterally optimal for the home government to utilize its trade policies in a way that does *not* fully correct the international hold-up problem.

We can thus conclude that, when only H intervenes, international efficiency is not achieved. Instead, there are now *two* sources of international inefficiency that arise: an inefficiently low input trade volume that results from the continued existence of the international hold-up problem; and distortions in the final good market that arise as a result of the home-country's attempts to extract bargaining surplus from foreign suppliers. We may thus state:

Proposition 3 *In the Benchmark Model, when only H intervenes with trade policy, its unilaterally optimal policy choices lead to (i) an inefficiently low volume of input trade ($\hat{x} < x^E$), and (ii) an inefficiently low local price for the final good in H's market.*

Proposition 3 stands in marked contrast to Proposition 2, and reflects a simple point. To the extent that home-country producers share part of the surplus from production with foreign suppliers (as is the case in our Benchmark Model), the unilateral incentives of the home-country government to intervene with trade policy to mitigate the international hold-up problem will be muted by the fact that foreign suppliers enjoy some of the benefits of this intervention. In this environment, the home-country's unilateral intervention must be concerned as well with capturing foreign surplus, and therefore the home country cannot be counted on to solve the international hold-up problem on its own. Moreover, the home-country's attempts to extract bargaining surplus from foreign suppliers will spill over into the final good market as well, and introduce additional distortions there.

3.2 Foreign Intervention and Nash Policy Choices

We turn next to consider the unilateral incentives of the foreign government to intervene with a trade tax τ_x^F (as before, in a prior *stage* 0). We hence modify *stage* 0 as follows:

stage 0. The home government H selects a trade tax τ_1^H on the final good 1, and a trade tax τ_x^H on the imported input x; simultaneously, the foreign government F selects a trade tax τ_x^F on the exported input x.[21]

We start by considering F's incentive to intervene facing a given H policy pair $\left(\tau_1^H, \tau_x^H\right)$. In this case, the input supplier in F now has a payoff of

$$\pi^F = \frac{1}{2}(1+\tau_1^H)y\left(\hat{x}\right) - (1+\frac{1}{2}\tau_x^H+\frac{1}{2}\tau_x^F)\hat{x},$$

with $\hat{x}$ defined by

$$\frac{1}{2}(1+\tau_1^H)y'\left(\hat{x}\right) = 1+\frac{1}{2}\tau_x^H+\frac{1}{2}\tau_x^F. \tag{15}$$

Foreign welfare is then given by the sum of foreign consumer surplus, profits and tax revenue:

$$W^F = CS^F(1) + \frac{1}{2}\left(1+\tau_1^H\right)y\left(\hat{x}\right) - (1+\frac{1}{2}\tau_x^H+\frac{1}{2}\tau_x^F)\hat{x} + \tau_x^F\hat{x}.$$

The optimal choice of τ_x^F, which we denote by $\hat{\tau}_x^F$, hence must satisfy the first-order condition

$$\frac{\partial W^F}{\partial \tau_x^F} = 0 = \frac{1}{2}\hat{x} + [\frac{1}{2}\left(1+\tau_1^H\right)y'\left(\hat{x}\right) - 1 - \frac{1}{2}\tau_x^H + \frac{1}{2}\tau_x^F)]\frac{\partial \hat{x}}{\partial \tau_x^F}. \tag{16}$$

Recalling that $\partial\hat{x}/\partial\tau_x^F < 0$, the first-order condition in (16) together with (15) immediately implies that

$$\tau_x^F = -\frac{1}{2}\frac{\hat{x}}{\partial\hat{x}/\partial\tau_x^F} > 0, \tag{17}$$

and hence, the foreign country finds it optimal to set an export tax on the intermediate input.

The logic behind this result can be understood as follows. First, why doesn't F's government offer an export subsidy to increase exports of x and help address the hold-up problem? The reason is that the level of x is already chosen by the foreign supplier to maximize foreign profits, and so there is no gain to the foreign country from manipulating this choice with export-sector intervention. And second, foreign suppliers do not bear the full cost of the increase in the marginal

[21] As in the case of second-best policies, we could allow for foreign taxes on trade in the final good 1, but these have no effect on the hold-up problem and will thus never be used as a part of an optimal set of policies.

cost of production associated with an export tax, because they have less than full bargaining power in their negotiations with final-good producers. Hence, the foreign government is able to pass part of the cost of the export tax on to the home country while keeping the entire benefit from it (in the form of tax revenue). As a result, the optimal export tax is positive.

How will the home country respond to the setting of an export tax by F? In order to derive the Nash policy choices of the home country, we next solve for the final-good tax τ_1^H and the input tax τ_x^H that maximize home welfare for a given foreign policy choice τ_x^F. This pair will thus maximize

$$W^H = CS(p_1^H) + \frac{1}{2}\left(\left(1+\tau_1^H\right) y\left(\hat{x}\right) - \tau_x^H \hat{x} - \tau_x^F \hat{x}\right) + \tau_1^H [D_1(p_1^H) - y\left(\hat{x}\right)] + \tau_x^H \hat{x}$$

subject to $\hat{x}$ being given by (15). Manipulating the first-order conditions and replacing τ_x^F with $\hat{\tau}_x^{FN} \equiv -\left(1/2\right)\hat{x}/\left(\partial \hat{x}/\partial \tau_x^F\right)$ yields the following conditions defining the home Nash pair $\left(\hat{\tau}_1^{HN}, \hat{\tau}_x^{HN}\right)$:

$$\begin{aligned}
\hat{\tau}_1^{HN} &= -\frac{\frac{1}{2}\hat{x}\left[\frac{y(\hat{x})}{\hat{x}} - y'\left(\hat{x}\right)\right]}{|\partial D_1/\partial p_1^H|}, \text{ and} \\
\hat{\tau}_x^{HN} &= -(1-\hat{\tau}_1^{HN})y'\left(\hat{x}\right) - \frac{\hat{x}}{\partial \hat{x}/\partial \tau_x^H} + \hat{\tau}_x^{FN}.
\end{aligned}$$

The first equation implies that $\hat{\tau}_1^{HN}$ is again negative, while the second indicates that the sign of $\hat{\tau}_x^{HN}$ is indeterminate. This parallels the results we obtain in the case without foreign retaliation as recorded in (11), and the intuition is the same as that outlined above. The only difference is the additional term $\hat{\tau}_x^{FN} > 0$ in the second equation, which other things equal leads to overall higher input taxes. Combining the above expressions for $\hat{\tau}_1^{HN}$ and $\hat{\tau}_x^{HN}$ with equations (15) and (17) we further obtain:

$$y'\left(\hat{x}\right) = 1 - \frac{\hat{x}}{\partial \hat{x}/\partial \tau_x^H}. \tag{18}$$

It is then clear that the Nash equilibrium involves suboptimal trade in intermediate inputs, $\hat{x} < x^E$. In sum, we have shown that:

Proposition 4 *In the Nash equilibrium of the Benchmark Model, F maintains free trade in the final good and taxes the exports of the input, while H intervenes in both the final-good and input markets, resulting in (i) an inefficiently low volume of input trade ($\hat{x} < x^E$), and (ii) an inefficiently low local price for the final good in H's market.*

4 The Role and Design of a Trade Agreement

In previous sections, we explore the logic of trade policy intervention in the presence of offshoring. In this section we consider the role and design of a trade agreement. To facilitate the discussion, it is useful to first recast the conditions for efficient and Nash policies in terms of the local and international prices that these policies induce. As we demonstrate, this helps to interpret the problem that a trade agreement can solve in this setting, and it also helps to identify the design features of a trade agreement that could help governments achieve the efficiency frontier.

To this end, we begin by defining the *international price* of the input x, which we denote by p_x^*. In words, p_x^* is the (untaxed) price negotiated in *stage* 3 for the exchange of inputs between the foreign supplier and the home producer. It is easy to see that in the Benchmark Model this price is given by $p_x^* = \pi^F/\hat{x} + (1+\tau_x^F)$, which can be written as

$$p_x^* = p_x^*(\tau_1^H, \tau_x^H, \tau_x^F) \equiv \frac{1}{2}(1+\tau_1^H)\frac{y\left(\hat{x}(\tau_1^H, \tau_x^H + \tau_x^F)\right)}{\hat{x}(\tau_1^H, \tau_x^H + \tau_x^F)} - \frac{1}{2}\left(\tau_x^H - \tau_x^F\right), \tag{19}$$

where $\hat{x}(\tau_1^H, \tau_x^H + \tau_x^F)$ is defined by (15). Given that the world price of the final good 1 is fixed by assumption, the international price p_x^* plays the role of the terms of trade between the Home and Foreign country in the Benchmark Model. Notice that (19) implies

$$\frac{\partial p_x^*}{\partial \tau_x^F} - \frac{\partial p_x^*}{\partial \tau_x^H} = 1. \tag{20}$$

Using (19) and (15), it can also be derived that

$$\frac{\partial p_x^*}{\partial \tau_x^H} = \frac{1}{-2\hat{x}y''}\left(\left[\frac{y(\hat{x})}{\hat{x}} - y'(\hat{x})\right] + \hat{x}y''(\hat{x})\right) < 0,$$

where the inequality follows from condition (2). Hence, a rise in τ_x^H improves the Home terms of trade. Moreover, it is easy to show that $\frac{\partial p_x^*}{\partial \tau_x^H}$ must be greater than -1, and therefore (20) implies that $\frac{\partial p_x^*}{\partial \tau_x^F}$ must be greater than 0. Hence, a rise in τ_x^F improves the Foreign terms of trade.

Next, recalling that $\tau_x \equiv \tau_x^H + \tau_x^F$ and noting that we can then write $\hat{x}(\tau_1^H, \tau_x)$, we define the Home-country price of the input x by

$$p_x^H \equiv p_x^* + \tau_x^H = \frac{1}{2}(1+\tau_1^H)\frac{y\left(\hat{x}(\tau_1^H, \tau_x)\right)}{\hat{x}(\tau_1^H, \tau_x)} + \frac{1}{2}\tau_x \equiv p_x^H(\tau_1^H, \tau_x).$$

Similarly, we define the Foreign-country price of the input x by

$$p_x^F \equiv p_x^* - \tau_x^F = \frac{1}{2}(1+\tau_1^H)\frac{y\left(\hat{x}(\tau_1^H, \tau_x)\right)}{\hat{x}(\tau_1^H, \tau_x)} - \frac{1}{2}\tau_x \equiv p_x^F(\tau_1^H, \tau_x).$$

Finally, notice that $p_x^H - p_x^F = \tau_x$, and recall that $p_1^H = 1 + \tau_1^H$. This implies that we may express $\hat{x}$ equivalently as a function of local Home and Foreign prices: $\hat{x}(\tau_1^H, \tau_x) = \bar{x}(p_1^H, p_x^H - p_x^F)$. Below we will continue to make use of the function $\hat{x}(\tau_1^H, \tau_x)$, but it will sometimes be convenient to use the equivalent function $\bar{x}(p_1^H, p_x^H - p_x^F)$.

With these definitions, we are now ready to express Home and Foreign welfare as functions of local and international prices. In particular, letting $\bar{x}(\cdot)$ denote $\bar{x}(p_1^H, p_x^H - p_x^F)$ for notational ease, we may write

$$\begin{aligned} W^H &= CS(p_1^H) + [p_1^H y(\bar{x}(\cdot)) - p_x^H \bar{x}(\cdot)] + (p_1^H - 1)[D(p_1^H) - y(\bar{x}(\cdot))] + (p_x^H - p_x^*)\bar{x}(\cdot) \\ &\equiv \bar{W}^H \left(p_1^H(\tau_1^H),\ p_x^H(\tau_1^H, \tau_x), p_x^F(\tau_1^H, \tau_x),\ p_x^*(\tau_1^H, \tau_x^H, \tau_x^F)\right), \end{aligned} \tag{21}$$

and

$$\begin{aligned} W^F &= CS(1) + \bar{x}(\cdot)[p_x^F - 1] + (p_x^* - p_x^F)\bar{x}(\cdot) = CS(1) + (p_x^* - 1)\bar{x}(\cdot) \\ &\equiv \bar{W}^F \left(p_1^H(\tau_1^H),\ p_x^H(\tau_1^H, \tau_x), p_x^F(\tau_1^H, \tau_x),\ p_x^*(\tau_1^H, \tau_x^H, \tau_x^F)\right). \end{aligned} \tag{22}$$

Here and throughout this section, we use $\bar{W}^j$ to represent the objectives of government j when expressed as a function *only* of prices.

Notice that, with subscripts on the welfare functions denoting partial derivatives, expressions (21) and (22) imply

$$\bar{W}^H_{p_x^*} = -\hat{x} \text{ and } \bar{W}^F_{p_x^*} = \hat{x}, \tag{23}$$

and so $\bar{W}^H_{p_x^*} + \bar{W}^F_{p_x^*} = 0$. This reflects the fact that the income effect of the terms-of-trade change embodied in the rise of p_x^* – holding local prices fixed – is given simply by the trade volume ($\hat{x}$), and amounts to a pure (inframarginal) transfer of rents from the home country to the foreign country. This property is also reflected in the fact that the sum of Home and Foreign welfare is independent of p_x^*. In particular, we may write:

$$\begin{aligned} \bar{W}^W &\equiv \bar{W}^H + \bar{W}^F \\ &= CS(p_1^H) + [p_1^H y(\bar{x}(\cdot)) - p_x^H \bar{x}(\cdot)] + (p_1^H - 1)[D(p_1^H) - y(\bar{x}(\cdot))] \\ &\quad + CS(1) + (p_x^H - 1)\bar{x}(\cdot) \\ &\equiv \bar{W}^W(p_1^H(\tau_1^H),\ p_x^H(\tau_1^H, \tau_x), p_x^F(\tau_1^H, \tau_x)). \end{aligned}$$

An implication is that efficiency imposes conditions only on τ_1^H and τ_x, confirming the analogous finding reported in the previous section.

Using the welfare expressions given in (21) and (22) and the prices defined above, we may now express the conditions that efficient policies τ_1^{HE} and τ_x^E must satisfy:

$$\begin{aligned} \bar{W}^W_{p_x^H}\frac{\partial p_x^H}{\partial \tau_x} + \bar{W}^W_{p_x^F}\frac{\partial p_x^F}{\partial \tau_x} &= 0, \text{ and} \\ \bar{W}^W_{p_1^H} + \bar{W}^W_{p_x^H}\frac{\partial p_x^H}{\partial \tau_1^H} + \bar{W}^W_{p_x^F}\frac{\partial p_x^F}{\partial \tau_1^H} &= 0, \end{aligned} \tag{24}$$

where we have used the fact that $\frac{\partial p_1^H}{\partial \tau_1^H} = 1$. By solving the top expression in (24) for $\bar{W}^W_{p_x^F}$ and using the resulting expression to eliminate $\bar{W}^W_{p_x^F}$ from the bottom expression in (24), and by observing that changes in τ_x^H and τ_1^H that hold fixed $p_x^*(\tau_1^H, \tau_x^H, \tau_x^F)$ must hold fixed as well the foreign local price $p_x^F(\tau_1^H, \tau_x) = p_x^* - \tau_x^F$ given that τ_x^F is unchanged, and hence are defined by $\frac{d\tau_x^H}{d\tau_1^H}|_{dp_x^*=0} = \frac{-\partial p_x^F/\partial \tau_1^H}{\partial p_x^F/\partial \tau_x}$, the conditions for efficiency in (24) can be rewritten in an equivalent form that is convenient for later comparison:

$$\begin{aligned}\bar{W}^W_{p^H_x}\frac{\partial p^H_x}{\partial \tau_x}+\bar{W}^W_{p^F_x}\frac{\partial p^F_x}{\partial \tau_x} &= 0, \text{ and} \\ \bar{W}^W_{p^H_1}+\bar{W}^W_{p^H_x}\left(\frac{\partial p^H_x}{\partial \tau^H_1}+\frac{\partial p^H_x}{\partial \tau_x}\left.\frac{d\tau^H_x}{d\tau^H_1}\right|_{dp^*_x=0}\right) &= 0,\end{aligned} \tag{25}$$

where we have used $\frac{d\tau_x}{d\tau^H_x}=1$. The top condition says that at efficient policies, a small change in τ_x must have no first-order impact on world welfare. The bottom condition says that in addition, small changes in τ^H_1 and τ^H_x that hold fixed p^*_x and hence p^F_x must have no first-order impact on world welfare either.

It is interesting to pause here and consider further the bottom efficiency condition in (25). In a setting where international prices are determined by market clearing, changes in home-country policies that leave equilibrium international prices unchanged must also leave international trade volumes unchanged, and hence must leave foreign-country welfare unchanged, a feature that supplies the basis for the well-known Kemp-Wan Theorem (Kemp and Wan, 1976) and that is emphasized as well in Bagwell and Staiger (2001a) as having important implications for the design of trade agreements. But here, with international prices determined by bilateral bargaining, this feature does not hold. In fact, using (15) and (19), it can be shown that

$$\frac{\partial \hat{x}(\tau^H_1,\tau_x)}{\partial \tau^H_1}+\frac{\partial \hat{x}(\tau^H_1,\tau_x)}{\partial \tau_x}\left.\frac{d\tau^H_x}{d\tau^H_1}\right|_{dp^*_x=0}=\frac{\left[\frac{y(\hat{x})}{\hat{x}}-y'(\hat{x})\right]\hat{x}}{p^H_1\left(\left[\frac{y(\hat{x})}{\hat{x}}-y'(\hat{x})\right]+\hat{x}y''\right)}<0, \tag{26}$$

where the inequality again follows from (2). Evidently, an increase in τ^H_1 that is accompanied by a change in τ^H_x which prevents p^*_x from changing must *reduce* the equilibrium volume of input trade $\hat{x}$, and will therefore in general have impacts on foreign welfare as well as home welfare. For this reason, at the efficient policies, the impacts of these tariff changes on both Home and Foreign welfare must be considered, as is reflected in the bottom efficiency condition in (25).

Consider next the Nash policy choices, τ^{HN}_1, τ^{HN}_x and τ^{FN}_x. Using the fact that $\frac{d\tau_x}{d\tau^H_x}=1=\frac{d\tau_x}{d\tau^F_x}$, these policies are defined by the solutions to the three first-order conditions

$$\begin{aligned}\bar{W}^H_{p^H_1}+\bar{W}^H_{p^H_x}\frac{\partial p^H_x}{\partial \tau^H_1}+\bar{W}^H_{p^F_x}\frac{\partial p^F_x}{\partial \tau^H_1}+\bar{W}^H_{p^*_x}\frac{\partial p^*_x}{\partial \tau^H_1} &= 0, \\ \bar{W}^H_{p^H_x}\frac{\partial p^H_x}{\partial \tau_x}+\bar{W}^H_{p^F_x}\frac{\partial p^F_x}{\partial \tau_x}+\bar{W}^H_{p^*_x}\frac{\partial p^*_x}{\partial \tau^H_x} &= 0, \text{ and} \\ \bar{W}^F_{p^H_x}\frac{\partial p^H_x}{\partial \tau_x}+\bar{W}^F_{p^F_x}\frac{\partial p^F_x}{\partial \tau_x}+\bar{W}^F_{p^*_x}\frac{\partial p^*_x}{\partial \tau^F_x} &= 0.\end{aligned} \tag{27}$$

To confirm that Nash policies are inefficient, we add together the middle and bottom expressions in (27) and use (20) and (23) to derive a first implication of Nash policies:

$$\bar{W}^W_{p^H_x}\frac{\partial p^H_x}{\partial \tau_x}+\bar{W}^W_{p^F_x}\frac{\partial p^F_x}{\partial \tau_x}=-\hat{x}^N, \tag{28}$$

where $\hat{x}^N \equiv \hat{x}(\tau_1^{HN}, \tau_x^N)$ with $\tau_x^N \equiv \tau_x^{HN}+\tau_x^{FN}$. And next, proceeding as before, we solve the middle expression in (27) for $\bar{W}^W_{p^F_x}$ and use the resulting expression to eliminate $\bar{W}^W_{p^F_x}$ from the top expression in (27) to derive a second implication of the Nash policies:

$$\bar{W}^H_{p^H_1}+\bar{W}^H_{p^H_x}\left(\frac{\partial p^H_x}{\partial \tau^H_1}+\frac{\partial p^H_x}{\partial \tau_x}\left.\frac{d\tau^H_x}{d\tau^H_1}\right|_{dp^*_x=0}\right)=0. \tag{29}$$

It is direct from (28) that the top efficiency condition in (25) is not satisfied at the Nash policies, and that world welfare could be increased by reducing τ_x below the Nash level τ_x^N. Moreover, it is clear from (29) that the bottom efficiency condition in (25) is satisfied at Nash policies if and only if at Nash policies it is also true that

$$\bar{W}^F_{p^H_1}+\bar{W}^F_{p^H_x}\left(\frac{\partial p^H_x}{\partial \tau^H_1}+\frac{\partial p^H_x}{\partial \tau_x}\left.\frac{d\tau^H_x}{d\tau^H_1}\right|_{dp^*_x=0}\right)=0. \tag{30}$$

Intuitively, as we have emphasized above, the changes in τ_1^H and τ_x^H that hold fixed p_x^* as described in the bottom efficiency condition of (25) will in general impact both Home and Foreign welfare; but as the Nash condition (29) indicates, when the home government selects its Nash levels of τ_1^H and τ_x^H it is of course sensitive to how such adjustments impact its own welfare, but it is not sensitive to how such adjustments impact the welfare of the foreign country. Hence, these Nash policy choices will only be efficient if (30) holds so that there is, in fact, no impact on the welfare of the foreign country of changes in τ_1^H and τ_x^H that hold fixed p_x^* when these impacts are evaluated at Nash policies. But it can be confirmed that, when evaluated at Nash policies, (30) does not hold and that the expression on the left-hand-side of (30) is strictly negative, which is to say that at the Nash policies the foreign country is hurt by the reduction in $\hat{x}$ implied by these policy adjustments; and so it is apparent that the bottom efficiency condition in (25) is not satisfied at the Nash policies either.

Of course, this discussion simply confirms what we have already established in the previous section, that Nash policies are inefficient. But as we observed above and next demonstrate, developing an understanding for the reason for the inefficiencies in the Nash equilibrium is facilitated by expressing the Nash and efficiency conditions in this form.

4.1 Identifying the Problem

We now ask whether the inefficiency of Nash policies can be attributed to terms-of-trade manipulation. To this end, we follow Bagwell and Staiger (1999, 2009a,b) and define *politically optimal* tariffs as those tariffs that would hypothetically be chosen by governments unilaterally if they did not value the pure international rent-shifting associated with the terms-of-trade movements induced by their unilateral tariff choices. Specifically, we suppose that the home government acts as if $\bar{W}^H_{p^*_x} \equiv 0$ when choosing its politically optimal tariffs, while the foreign government acts as if $\bar{W}^F_{p^*_x} \equiv 0$ when choosing its politically optimal tariff. We therefore define politically optimal tariffs, which we denote by τ_1^{HPO}, τ_x^{HPO} and τ_x^{FPO}, as those tariffs that satisfy the three conditions

$$\begin{aligned}
\bar{W}^H_{p_1^H} + \bar{W}^H_{p_x^H}\frac{\partial p_x^H}{\partial \tau_1^H} + \bar{W}^H_{p_x^F}\frac{\partial p_x^F}{\partial \tau_1^H} &= 0, \\
\bar{W}^H_{p_x^H}\frac{\partial p_x^H}{\partial \tau_x} + \bar{W}^H_{p_x^F}\frac{\partial p_x^F}{\partial \tau_x} &= 0, \text{ and} \\
\bar{W}^F_{p_x^H}\frac{\partial p_x^H}{\partial \tau_x} + \bar{W}^F_{p_x^F}\frac{\partial p_x^F}{\partial \tau_x} &= 0.
\end{aligned} \tag{31}$$

Having defined politically optimal tariffs in this way, we may then ask whether politically optimal tariffs are efficient, and thereby determine whether the Nash inefficiencies identified above can be given a terms-of-trade interpretation, according to which the fundamental problem faced by governments in designing their trade agreement is to find a way to eliminate terms-of-trade manipulation.

To assess the efficiency properties of politically optimal tariffs, we add together the middle and bottom expressions in (31) to derive a first implication of politically optimal policies:

$$\bar{W}^W_{p_x^H}\frac{\partial p_x^H}{\partial \tau_x} + \bar{W}^W_{p_x^F}\frac{\partial p_x^F}{\partial \tau_x} = 0. \tag{32}$$

And we solve the middle expression in (31) for $\bar{W}^W_{p_x^F}$ and use the resulting expression to eliminate $\bar{W}^W_{p_x^F}$ from the top expression in (31) to derive a second implication of politically optimal policies:

$$\bar{W}^H_{p_1^H} + \bar{W}^H_{p_x^H}\left(\frac{\partial p_x^H}{\partial \tau_1^H} + \frac{\partial p_x^H}{\partial \tau_x}\left.\frac{d\tau_x^H}{d\tau_1^H}\right|_{dp_x^*=0}\right) = 0. \tag{33}$$

It is direct from (32) that the top efficiency condition in (25) is satisfied at politically optimal tariffs. Moreover, at politically optimal policies it can be confirmed that

$$\bar{W}^F_{p_1^H} + \bar{W}^F_{p_x^H}\left(\frac{\partial p_x^H}{\partial \tau_1^H} + \frac{\partial p_x^H}{\partial \tau_x}\left.\frac{d\tau_x^H}{d\tau_1^H}\right|_{dp_x^*=0}\right) = 0 \tag{34}$$

and (33) and (34) together imply that the bottom efficiency condition in (25) is satisfied at politically optimal tariffs as well. Hence, politically optimal policies are efficient in this setting, and the problem for a trade agreement to solve can be given a terms-of-trade interpretation.

Before considering what this implies for the design of trade agreements in the presence of offshoring, it is instructive to consider further why it is that politically optimal tariffs are efficient. As well, it is interesting to consider how it could be that the foreign country is unaffected by the changes in τ_1^H and τ_x^H described by (33) when these changes are evaluated at the politically optimal tariffs, as (34) reflects. An understanding of these points can be developed by examining the specific form that the politically optimal policies take.

Using the explicit welfare expressions contained in (21) and (22) and explicitly evaluating derivatives, the conditions for politically optimal policies in (31) reduce to:

$$\begin{aligned}(p_1^H - 1)D' - \frac{y'(\hat{x})}{p_1^H y''(\hat{x})}(1-p_x^*) &= 0, \\ y'(\hat{x}) - p_x^* &= 0, \text{ and} \\ 1 - p_x^* &= 0.\end{aligned} \tag{35}$$

By examining the three conditions in (35), it is possible to understand at a more intuitive level why terms-of-trade manipulation is the source of the inefficiency of Nash policy choices in the Benchmark Model.

Consider first the determination of τ_x^{FPO}. As the bottom expression in (35) indicates, at the political optimum the foreign government selects a tariff that equates the international price of its exported input x with the (unitary) marginal cost of production of x: intuitively, the foreign government's interests are identical to those of a monopolist with regard to its exports of x; and so, when the foreign government hypothetically does not value the pure (inframarginal) rents generated by changes in the international price of its exports, as in the thought experiment that defines the politically optimal policies, it prefers to drive export volume to the competitive level where price is equated with marginal cost. With the choice of τ_x^{FPO} guaranteeing $p_x^* = 1$, the middle expression of (35) then implies the home government's selection of τ_x^{HPO} ensures that $y'(\hat{x}) = 1$, and hence by (3) that $\hat{x}$ is brought to its efficient level: intuitively, the home government's interests are identical to those of a monopsonist with regard to its imports of x; and so, when the home government hypothetically does not value the pure (inframarginal) rents generated by changes in the international price of its imported input, it perceives the international price p_x^* to be the marginal cost of an additional imported unit and drives imports of x to the level at which marginal revenue valued at world prices ($y'(\hat{x})$) is equal to perceived marginal cost.[22] Finally, in light of the bottom expression in (35), it is clear from the top expression in (35) that the home government has no incentive to distort its final-good price p_1^H away from the world price of 1 when it is not concerned with terms-of-trade manipulation, and so the politically optimal final-good tariff for the home country is free trade (i.e., $\tau_1^{HPO} = 0$).

We can now also understand how it is that the political optimum can achieve efficiency despite the fact that, as (33) confirms, the home government chooses its politically optimal policies without regard to the impact on Foreign welfare of changes in τ_1^H and τ_x^H that hold fixed p_x^* but reduce the equilibrium volume of input trade $\hat{x}$ (as (26) indicates). The reason is that the foreign country is itself choosing τ_x^F in the political optimum so that, at the level of p_x^* implied by politically optimal

[22] While it is clear from the middle condition of (35) that $\tau_x^{PO} \equiv \tau_x^{HPO} + \tau_x^{FPO} = \tau_x^E$, a remaining question is what determines the precise levels of τ_x^{HPO} and τ_x^{FPO}. It can be shown that τ_x^{HPO} and τ_x^{FPO} are unique and that both lie in the open interval $(-1, 0)$, with the precise levels determined by the curvature properties of $y(x)$.

policies, price is equated to marginal cost and the foreign country is therefore, in fact, indifferent to small changes in the export volume $\hat{x}$.

We summarize these results with:

Proposition 5 *In the Benchmark Model, the problem for a trade agreement to solve can be given a terms-of-trade interpretation.*

4.2 Implications for the Design of Trade Agreements

If terms-of-trade manipulation is the problem that drives inefficiencies in the Nash equilibrium of the Benchmark Model, then it seems natural to expect that the same features of the GATT/WTO that have been shown to work well in principle to serve governments in their efforts to escape from a terms-of-trade driven Prisoners' Dilemma in other settings (see, for example, Bagwell and Staiger, 1999, 2001a, and most recently 2009a,b) should carry over in the presence of offshoring. In fact, as we next demonstrate, the GATT/WTO principle of reciprocity continues to exhibit attractive features in the presence of offshoring. But there is one central feature of the GATT/WTO system whose desirable properties do not carry over to a setting such as the Benchmark Model in which offshoring is present: this feature is the GATT/WTO focus on "market access," according to which negotiations are narrowly focused on the tariffs that are applied directly to the trade flows in question, and where the concept of "nullification or impairment" is then utilized to protect the implied access from subsequent unilateral adjustments that a government might wish to make to its wider portfolio of policies.

To establish these points, we begin by deriving the implications of reciprocity in the Benchmark Model. Formally, we follow Bagwell and Staiger (1999, 2001b) and define tariff changes that conform to reciprocity as those that bring about equal changes in the volume of each country's imports and exports when valued at existing world prices. Accounting for trade in the numeraire good 0 and the final good 1 as well as the input x, and letting a superscript "A" denote original trade tax levels and a superscript "B" denote new trade tax levels, it is direct to establish that tariff changes conforming to reciprocity must satisfy[23]

$$[p_x^*(\tau_1^{HB}, \tau_x^{HB}, \tau_x^{FB}) - p_x^*(\tau_1^{HA}, \tau_x^{HA}, \tau_x^{FA})]\hat{x}(\tau_1^{HB}, \tau_x^{B}) = 0. \tag{36}$$

Evidently, as (36) indicates, tariff changes that conform to reciprocity (and that do not eliminate trade in the input x) must leave the international price p_x^* unchanged.

Armed with this result, it is straightforward to establish that the two attractive features of reciprocity that have been highlighted in other settings also apply here: namely, that beginning from Nash policies, each country can enjoy mutual gains from at least a small amount of tariff liberalization that conforms to reciprocity; and that beginning from efficient politically optimal

[23] The steps to derive (36) use the fact that the world price of good 1 is fixed by our small country assumption and employ the balanced trade condition that must hold at the original and the new world prices, and are identical to those described in note 19 of Bagwell and Staiger (2001b).

policies, no country would wish to deviate from these policies if the country understood that such a deviation would be met by reciprocal adjustments in the trade policies of its trading partners. These features simply reflect the fact that, if terms-of-trade manipulation is the problem, then a rule that prevents trade policy changes from altering the terms of trade is bound to have some attractive properties. Hence, in the Benchmark Model, as in other models where the Nash inefficiency can be traced to terms-of-trade manipulation, the principle of reciprocity serves to "undo" the terms-of-trade driven restrictions in trade volume that occur when governments pursue unilateral trade policies.

We next consider the GATT/WTO focus on market access. As Bagwell and Staiger (2001a) have demonstrated, in a wide variety of settings where terms-of-trade manipulation is the problem for a trade agreement to solve, it is possible to reach the international efficiency frontier with trade negotiations that are structured so as to achieve efficient trade volumes based on tariff commitments and the market access that these commitments imply, leaving each country free to choose the best way to deliver this access with its preferred mix of trade and domestic policies. The access commitment implied by negotiated tariff levels is protected in the GATT/WTO by various nullification-or-impairment provisions, and as Bagwell and Staiger show this commitment can be interpreted formally as a promise not to alter the terms of trade with subsequent unilateral policy adjustments. While Bagwell and Staiger consider tariffs and domestic standards, the analogous result in our model of offshoring would be that governments can achieve efficient policy combinations by negotiating over τ_x^H and τ_x^F alone and then permitting the home country to make unilateral adjustments to τ_1^H and τ_x^H provided that these adjustments preserve the terms of trade p_x^*.

Importantly, the settings considered by Bagwell and Staiger (2001a) all share the property that international prices are determined by market-clearing conditions, and this leads to a feature that is key for the result: as Bagwell and Staiger emphasize, policy adjustments by one country that do not alter its terms of trade do not alter trade volumes either, and hence cannot effect the country's trading partners. And it is because of this feature that a country acting unilaterally can be expected to make internationally efficient policy adjustments when it is held to its market-access commitments. As we have emphasized above, however, this feature does not hold in the Benchmark Model, because the international price p_x^* is determined by bilateral bargaining between the Foreign seller and the Home buyer, not by an anonymous market clearing condition. Rather, as (26) confirms, in the Benchmark Model adjustments in the Home policies τ_1^H and τ_x^H that preserve p_x^* nevertheless alter the volume of input trade $\hat{x}$, and so as a general matter such adjustments will alter Foreign welfare and cannot be counted on to lead the home country to make efficient unilateral policy adjustments.[24] As a consequence, reaching the efficiency frontier generally requires

[24] Our model of offshoring differs from the typical model used in the trade-agreements literature in two ways: first, it emphasizes input trade; and second, the international price of the traded input is determined by bilateral bargaining. We establish in Appendix B, however, that a model of input trade in which the international price of the traded input is determined by a market clearing condition does not exhibit the novel properties that our offshoring model exhibits, which is why we can attribute these properties to the novel manner in which international prices are determined in our model. Note also that, in principle, there is still one point on the efficiency frontier that could be delivered under the described negotiation structure, and that is the political optimum. The reason is that, while

negotiations over τ_x^H and τ_x^F *and* τ_1^H in the Benchmark Model.

We summarize this discussion with the following:

Proposition 6 *In the presence of offshoring, the principle of reciprocity continues to "undo" the terms-of-trade driven restrictions in trade volume that occur when governments pursue unilateral trade policies, but to achieve efficiency policy outcomes negotiations must generally cover a wider set of policy instruments than would be required in the absence of offshoring.*

5 The Benchmark Model with Political Economy

We have thus far assumed that each country's government is benevolent and seeks to maximize the aggregate welfare of its residents. Both casual and formal evidence suggest, however, that it is more realistic to formulate a social welfare function that weights asymmetrically the welfare of different groups in society. The political economy literature has stressed the role of special interest groups in generating these biases in policy (Baron, 1994, Grossman and Helpman, 1996).

In this section, we extend the Benchmark Model to allow for government welfare functions that place a higher weight on producer welfare than on consumer welfare. In light of analogous results reported for example in Grossman and Helpman (1994) and Bagwell and Staiger (2002, Ch. 10), it might be expected that the introduction of political economy motives can eliminate unrealistic features of the Benchmark Model's policy predictions (e.g., convert import subsidies to import tariffs), and in Appendix A.2 we show that this is indeed the case. In the main text, we focus on a different point: when the foreign government is motivated by political economy concerns and wishes to redistribute surplus towards its input producers, a novel inefficiency can be identified in the Nash equilibrium which is not associated with terms-of-trade manipulation. Below we establish this point by showing that the political optimum is no longer efficient, and we illustrate and interpret the novel role for a trade agreement in this setting by characterizing policy adjustments from the politically optimal policies which result in Pareto improvements for the governments.

5.1 Introducing Political Economy

To represent political-economy motives, we implicitly assume that producers are in a better position to solve the "collective action" problem and hence can better coordinate their demands on the government. We also assume that the ownership of productive assets is highly concentrated, so that we can ignore the role of producers as consumers and as receivers of lump-sum tax rebates. In particular, we let:

the political optimum ensures that Foreign is indifferent to small changes in the volume of input trade as we have observed, under our assumption that x is produced at constant marginal cost Foreign is in fact indifferent at the political optimum to large changes in the volume of input trade as well; and so, under the politically optimal Foreign tariff, the reduced access to the Home market implied by the permitted unilateral policy adjustments that would be required to bring the Home final-good tariff to its politically optimal level would not "nullify-or-impair" the Foreign country (i.e., would not reduce Foreign welfare). We do not emphasize this point, however, because it is an artifact of our assumption that the Foreign industry supply curve of x is horizontal.

$$W^j = CS^j + \gamma^j \pi^j + \text{Trade Tax Revenue}^j, \qquad \text{with } \gamma^j \geq 1, \text{ for } j \in \{H, F\}, \tag{37}$$

where γ^j represents the weight that the government of country j places on the welfare of its producers, with political-economy motives present in country j if and only if $\gamma^j > 1$. Using (21), (22) and (37), the welfare of the home and foreign governments in the (politically augmented) Benchmark Model can be written as

$$\begin{aligned} W^H &= CS(p_1^H) + \gamma^H [p_1^H y(\bar{x}(\cdot)) - p_x^H \bar{x}(\cdot)] + (p_1^H - 1)[D(p_1^H) - y(\bar{x}(\cdot))] + (p_x^H - p_x^*)\bar{x}(\cdot) \\ &\equiv \tilde{W}^H \left(p_1^H(\tau_1^H),\ p_x^H(\tau_1^H, \tau_x), p_x^F(\tau_1^H, \tau_x),\ p_x^*(\tau_1^H, \tau_x^H, \tau_x^F) \right), \end{aligned} \tag{38}$$

and

$$\begin{aligned} W^F &= CS(1) + \gamma^F [p_x^F - 1]\bar{x}(\cdot) + (p_x^* - p_x^F)\bar{x}(\cdot) \\ &\equiv \tilde{W}^F \left(p_1^H(\tau_1^H),\ p_x^H(\tau_1^H, \tau_x), p_x^F(\tau_1^H, \tau_x),\ p_x^*(\tau_1^H, \tau_x^H, \tau_x^F) \right), \end{aligned} \tag{39}$$

while world welfare may now be written as

$$\begin{aligned} \tilde{W}^W &\equiv \tilde{W}^H + \tilde{W}^F \\ &= CS(p_1^H) + \gamma^H [p_1^H y(\bar{x}(\cdot)) - p_x^H \bar{x}(\cdot)] + (p_1^H - 1)[D(p_1^H) - y(\bar{x}(\cdot))] \\ &\quad + CS(1) + \gamma^F [p_x^F - 1]\bar{x}(\cdot) + (p_x^H - 1)\bar{x}(\cdot) \\ &\equiv \tilde{W}^W (p_1^H(\tau_1^H),\ p_x^H(\tau_1^H, \tau_x), p_x^F(\tau_1^H, \tau_x)). \end{aligned}$$

Comparing (38) with (21) and (39) with (22), it is immediately clear that all of the analysis from section 4 continues to apply when political economy motives are introduced into the Benchmark Model, with the role of $\bar{W}^H$, $\bar{W}^F$ and $\bar{W}^W$ now replaced by $\tilde{W}^H$, $\tilde{W}^F$ and $\tilde{W}^W$ respectively, and with $\tilde{W}^H$ and $\tilde{W}^F$ sharing the properties of $\bar{W}^H$ and $\bar{W}^F$ contained in (23). We may therefore proceed directly to evaluating the efficiency properties of the Nash and politically optimal policies in the presence of political economy.

5.2 Identifying the Problem for a Trade Agreement to Solve

To confirm that Nash policy choices are inefficient in the presence of political economy forces, we simply proceed as in section 4 and compare the implications of the Nash policies (restated here with $\bar{W}$'s replaced by $\tilde{W}$'s),

$$\tilde{W}^W_{p_x^H} \frac{\partial p_x^H}{\partial \tau_x} + \tilde{W}^W_{p_x^F} \frac{\partial p_x^F}{\partial \tau_x} = -\hat{x}^N,$$

and

$$\tilde{W}^H_{p_1^H} + \tilde{W}^H_{p_x^H} \left(\frac{\partial p_x^H}{\partial \tau_1^H} + \frac{\partial p_x^H}{\partial \tau_x} \left. \frac{d\tau_x^H}{d\tau_1^H} \right|_{dp_x^*=0} \right) = 0,$$

to the conditions for efficiency (restated here with $\bar{W}$'s replaced by $\tilde{W}$'s),

$$\begin{aligned} \tilde{W}^W_{p^H_x}\frac{\partial p^H_x}{\partial \tau_x}+\tilde{W}^W_{p^F_x}\frac{\partial p^F_x}{\partial \tau_x} &= 0, \text{ and} \\ \tilde{W}^W_{p^H_1}+\tilde{W}^W_{p^H_x}\left(\frac{\partial p^H_x}{\partial \tau^H_1}+\frac{\partial p^H_x}{\partial \tau_x}\left.\frac{d\tau^H_x}{d\tau^H_1}\right|_{dp^*_x=0}\right) &= 0. \end{aligned} \tag{40}$$

As before, it is direct that the top efficiency condition in (40) is not satisfied at the Nash policies, and that the bottom efficiency condition in (40) is satisfied at Nash policies if and only if at Nash policies it is also true that

$$\tilde{W}^F_{p^H_1}+\tilde{W}^F_{p^H_x}\left(\frac{\partial p^H_x}{\partial \tau^H_1}+\frac{\partial p^H_x}{\partial \tau_x}\left.\frac{d\tau^H_x}{d\tau^H_1}\right|_{dp^*_x=0}\right)=0. \tag{41}$$

And it is straightforward to show that (41) is violated.

We may therefore conclude that Nash policies remain inefficient in the Benchmark Model when political economy considerations are introduced. This is not surprising, given that the terms-of-trade motive is still active when political economy motives are also present, and as we have already demonstrated terms-of-trade manipulation leads to inefficient policy choices. In our working paper (Antràs and Staiger, 2008) we also confirm a more specific finding: the particular nature of the inefficiencies associated with the Nash equilibrium as described in Proposition 4 still apply here. That is, even when governments are motivated by political-economy concerns, in the presence of offshoring the Nash volume of input trade and the local price for the final good in H's market are each inefficiently low (given the preferences of governments).

Next let us consider the efficiency properties of politically optimal policies. As we established in section 4, politically optimal policies imply the conditions (restated here with $\bar{W}$'s replaced by $\tilde{W}$'s)

$$\tilde{W}^W_{p^H_x}\frac{\partial p^H_x}{\partial \tau_x}+\tilde{W}^W_{p^F_x}\frac{\partial p^F_x}{\partial \tau_x}=0 \tag{42}$$

and

$$\tilde{W}^H_{p^H_1}+\tilde{W}^H_{p^H_x}\left(\frac{\partial p^H_x}{\partial \tau^H_1}+\frac{\partial p^H_x}{\partial \tau_x}\left.\frac{d\tau^H_x}{d\tau^H_1}\right|_{dp^*_x=0}\right)=0. \tag{43}$$

As before, it is direct that the top efficiency condition in (40) is satisfied at politically optimal tariffs. And as before, it is clear that the bottom efficiency condition in (40) is satisfied at politically optimal tariffs if and only if

$$\tilde{W}^F_{p^H_1}+\tilde{W}^F_{p^H_x}\left(\frac{\partial p^H_x}{\partial \tau^H_1}+\frac{\partial p^H_x}{\partial \tau_x}\left.\frac{d\tau^H_x}{d\tau^H_1}\right|_{dp^*_x=0}\right)=0. \tag{44}$$

But evaluated at the politically optimal tariffs, we have

$$\tilde{W}^F_{p^H_1}+\tilde{W}^F_{p^H_x}\left(\frac{\partial p^H_x}{\partial \tau^H_1}+\frac{\partial p^H_x}{\partial \tau_x}\left.\frac{d\tau^H_x}{d\tau^H_1}\right|_{dp^*_x=0}\right)=\frac{\left(\gamma^F-1\right)\hat{x}}{2}\left[\frac{y\left(\hat{x}\right)}{\hat{x}}-y'\left(\hat{x}\right)\right]>0. \tag{45}$$

Hence, in the presence of foreign political economy forces ($\gamma^F > 1$), politically optimal tariffs are not efficient, and so the Nash inefficiencies identified above cannot be given a terms-of-trade interpretation. Rather, as (43) and (45) indicate, beginning from politically optimal policies, a small increase in τ_1^H coupled with a change in τ_x^H that leaves p_x^* unchanged (and hence with τ_x^F fixed also leaves p_x^F unchanged) will lead to a second-order loss for Home (according to (43)) but results in a first-order gain for Foreign (according to (45)), and τ_x^H and τ_x^F can then be adjusted holding τ_x fixed so as to compensate Home for the second-order loss and still leave Foreign with a first-order gain from this maneuver.

It is instructive to consider further the nature of the additional Pareto gains that a trade agreement can generate in this setting beyond providing governments with an avenue of escape from a terms-of-trade driven Prisoners' Dilemma. To this end, recall from (26) that the equilibrium trade volume $\hat{x}$ must fall as a result of the policy adjustments described just above which improve upon the politically optimal policies. We may observe as well that at the political optimum, the politically motivated foreign government is offering an export subsidy to its input producers. Hence, with the foreign export subsidy held fixed at its politically optimal level $\tau_x^{FPO} < 0$, the changes in τ_1^H and τ_x^H described above induce budgetary savings for the foreign government in the amount of $\tau_x^{FPO}[\frac{\partial \hat{x}(\tau_1^H,\tau_x)}{\partial \tau_1^H} + \frac{\partial \hat{x}(\tau_1^H,\tau_x)}{\partial \tau_x}\frac{d\tau_x^H}{d\tau_1^H}|_{dp_x^*=0}]$, and these savings are evidently sufficiently valued by the foreign government at the political optimum to outweigh the cost to it of the reduction in income by the amount $\left(p_x^F - 1\right)[\frac{\partial \hat{x}(\tau_1^H,\tau_x)}{\partial \tau_1^H} + \frac{\partial \hat{x}(\tau_1^H,\tau_x)}{\partial \tau_x}\frac{d\tau_x^H}{d\tau_1^H}|_{dp_x^*=0}]$ that foreign input producers suffer. We may therefore conclude that the nature of the additional Pareto gains that we have characterized above amount to coordinating Home policies to help reduce the budgetary burden of the Foreign program of export promotion for the purposes of redistribution.

We summarize this result in

Proposition 7 *When the foreign government objectives include political economy considerations in the Benchmark Model, a trade agreement serves two roles: it provides governments with an avenue of escape from a terms-of-trade driven Prisoners' Dilemma; and it facilitates the setting of Home policies so as to reduce the budgetary burden of the Foreign program of export promotion for purposes of redistribution.*

Notice that, as is reflected in our discussion, it is the foreign political economy forces that prevent the politically optimal policies from being efficient. More generally, however, in the presence of symmetric home-supplier/foreign-producer relationships (which for example could be introduced into the Benchmark Model with the addition of a mirror-image second sector with the roles of Home and Foreign reversed), political economy forces in either country will interfere with the efficiency properties of the political optimum.

5.3 Implications for the Design of Trade Agreements

When offshoring is prevalent and governments pursue trade policies which maximize national income, the problem for a trade agreement to solve can be given a terms-of-trade interpretation, as

Proposition 5 suggests, and according to Proposition 6 the principle of reciprocity then continues to "undo" the terms-of-trade driven restrictions in trade volume that occur when governments pursue unilateral trade policies. However, as Proposition 6 indicates, to achieve efficient policy outcomes negotiations must generally cover a wider set of policy instruments than would be required in the absence of offshoring. Proposition 7 identifies a further distinction that arises when offshoring is prevalent: if governments are motivated by political economy concerns, a new inefficiency arises in the Nash equilibrium that cannot be given a terms-of-trade interpretation, and which instead relates to the budgetary consequences of export promotion.

Our Benchmark Model therefore formally identifies a separate "political externality" for a trade agreement to address when offshoring is prevalent. This externality bears a resemblance to that described in Ethier (2004).[25] And the identification of a political externality is in stark contrast to the predictions of the terms-of-trade theory, where the presence or absence of political economy motives has no impact on the underlying problem that a trade agreement must solve.

Finally, notice that Proposition 7 carries an important implication for the design of effective trade agreements in the presence of offshoring. Specifically, as Proposition 7 indicates, when offshoring and political economy forces are both present, the underlying problem that a trade agreement must address *varies* with the political preferences of member governments. As a consequence, under the view that governments operate in the presence of important political economy forces, the increasing prevalence of offshoring is likely to make it increasingly difficult for governments to rely on simple and general rules – such as reciprocity and non-discrimination – to help them solve their trade-related problems.

6 Sensitivity

In this section we consider the generality of our central findings to various alternative modeling assumptions. For simplicity, we return to a setting in which governments do not possess political economy motives.

6.1 General Bargaining Power

In the Benchmark Model we have assumed that all bargaining between home producers and foreign suppliers is characterized by symmetric Nash bargaining, with each party capturing one-half of the ex-post gains from trades. We have also assumed that the lack of an ex-post contractual agreement leaves both parties with no time to attempt to transact with alternative producers, and thus the outside options in the bargaining are equal to 0. We now explore the robustness of our results to the case in which bargaining power may be asymmetric and there exists a secondary market for

[25] As Ethier (2004, p. 305) puts it, " 'Political externalities,' by my definition, arise when policymakers in one country believe that their political status (whatever that might be specified to mean) is directly sensitive, to some degree, to actions by policymakers in another country." See also Bagwell and Staiger (2002, Chapter 2) for a further discussion of these ideas.

inputs at Home.[26] Our main conclusion is that the fundamental purposes of a trade agreement continue to be (a) helping governments solve the international hold-up problem between producers and suppliers, and (b) avoiding the policy distortions that would be introduced by attempts to extract bargaining surplus from foreign firms.

In order to explicitly derive the payoffs associated with the secondary market we now assume that good 1 comes in two types, a customized type T and a generic type G, and that consumer preferences are given by

$$U^j = c_0^j + u\left(c_{1_T}^j + \delta c_{1_G}^j\right), \quad 0 < \delta < 1. \tag{46}$$

Note that the preferences in (46) are such that consumers are willing to buy both types of good 1 only if the price of the generic relative to that of the customized type is equal to δ. This is analogous to consumers perceiving the two goods as perfect substitutes up to a quality shifter. By an appropriate choice of units, we can set the (fixed) price of customized inputs on world markets equal to 1, and that of generic inputs to δ.

The technology for producing final goods and intermediate inputs is as in our Benchmark Model. The only difference between the two types of good 1 is that the production of a generic good G uses an intermediate input x that is *not* customized to the producer's needs.

The game we consider is a straightforward extension to that in our Benchmark Model that incorporates a secondary market for inputs and generalized Nash bargaining. The sequence of events is as in our Benchmark Model, except that our previous *stages* 3 and 4 are divided into three stages as follows:

stage 3. Each producer-supplier pair bargains over the price of the intermediate input. We consider the generalized Nash bargaining solution with weights α and $(1-\alpha)$ for the home producer and foreign supplier, respectively, where $\alpha \in (0,1)$.

stage 4. A small number (formally, a measure-zero countable infinity) n of the bilateral pairs are exogenously dissolved and randomly rematched in a secondary market. They bargain again according the same generalized Nash bargaining solution as in *stage* 3. No further inputs can be produced; the amount produced in *stage* 2 is perceived as generic in the secondary market because it was tailored to another producer's specifications with probability one.

stage 5. Each producer in H imports x from its partner-supplier and produces the final good with the acquired x, and payments agreed in *stages* 3 and 4 are settled.

We focus directly on deriving Nash policy choices, assuming as before that the home and foreign governments select their respective tariffs simultaneously in a prior *stage* 0. Note that given the specification of the secondary market in *stage* 4, it is easy to see that the breakup of a single bargaining pair in *stage* 3 would result in each member of the pair being rematched with probability 1 with a random partner in *stage* 4, and therefore that *stage* 4 implies an outside option equal to

[26] In Antràs and Staiger (2008), we also considered the possibility that the secondary market for the foreign supplier is located in the foreign country.

$\alpha \left(\delta \left(1+\tau_1^H \right) y\left(x\right) -\left(\tau_x^H+\tau_x^F\right) x\right)$ for the final-good producer and $(1-\alpha)\left(\delta \left(1+\tau_1^H \right) y\left(x\right) -\left(\tau_x^H+\tau_x^F\right) x\right)$ for the supplier. These expressions are valid provided they are non-negative, and throughout this section we characterize results for the case where these non-negativity constraints are non-binding (though we showed in the Appendix of Antràs and Staiger, 2008, that our qualitative results carry through when these constraints bind).[27]

Following analogous steps as in previous sections, it is easy to see that generalized Nash bargaining in *stage* 3 will leave the final-good producer in H with a payoff equal to $\alpha \left(1+\tau_1^H \right) y\left(x\right) - \alpha \left(\tau_x^H+\tau_x^F\right) x$, with the supplier in F now receiving a *stage*-3 payoff of $(1-\alpha)(1+\tau_1^H)y(x) - (1-\alpha)(\tau_x^H+\tau_x^F)x$. This follows from the fact that the marginal cost of production of generic and customized inputs is the same, so there is no benefit in not customizing the input for the matched producer in *stage* 2.[28] As a result, we have that the *stage*-2 choice of $\hat{x}$ is now defined by

$$(1-\alpha)(1+\tau_1^H)y'(\hat{x}) = 1+(1-\alpha)\left(\tau_x^H+\tau_x^F\right), \tag{47}$$

and hence the *stage*-1 payoffs of the home and foreign firm are given by

$$\begin{aligned} \pi^H &= \alpha \left(1+\tau_1^H\right) y(\hat{x}) - \alpha\left(\tau_x^H+\tau_x^F\right)\hat{x}, \text{ and} \\ \pi^F &= (1-\alpha)(1+\tau_1^H)y(\hat{x}) - (1-\alpha)(\tau_x^H+\tau_x^F)\hat{x} - \hat{x}. \end{aligned}$$

Computing welfare in each country and solving for the first-order conditions that define the Nash policies $\hat{\tau}_1^{HN}$, $\hat{\tau}_x^{HN}$, and $\hat{\tau}_x^{FN}$, one can derive the following expressions, which are analogous to those in the Benchmark Model:[29]

$$\begin{aligned} \hat{\tau}_1^{HN} &= -\frac{(1-\alpha)\hat{x}\left[\frac{y(\hat{x})}{\hat{x}} - y'(\hat{x})\right]}{\left|\partial D_1/\partial p_1^H\right|}, \\ \hat{\tau}_x^{HN} &= -\frac{\left[\alpha-(1-\alpha)\hat{\tau}_1^{HN}\right]y'(\hat{x})}{(1-\alpha)} + \frac{\alpha\hat{\tau}_x^{FN}}{(1-\alpha)} - \frac{\hat{x}}{\partial\hat{x}/\partial\tau_x^H}, \text{ and} \\ \hat{\tau}_x^{FN} &= -\alpha\frac{\hat{x}}{\partial\hat{x}/\partial\tau_x^F}. \end{aligned}$$

By setting $\alpha = 1/2$, it is easily verified that these expressions coincide with those of the Benchmark Model above, where symmetric Nash bargaining and the absence of a secondary market for inputs was assumed. Considering the case of generalized Nash bargaining allows us to illustrate how the

[27] Beyond determining outside options, *stage* 4 plays no role in the model, and in particular only the customized type of good 1 will be produced with positive measure in equilibrium.

[28] Our framework emphasizes the role of customization in creating the lock-in effect at the heart of the holdup problem. As argued in section 2.1, however, the same lock-in effect could be generated by (ex-post) search frictions even in the absence of any customization. To see this, suppose that $\delta = 1$, so that generic and costumized inputs are perfect substitutes, but let search frictions lead to the formation of only κn pairs in *stage* 4, with $\kappa < 1$. It is then clear that the outside option for the final-good producer is now $\kappa\alpha\left(\left(1+\tau_1^H\right)y(x) - \left(\tau_x^H+\tau_x^F\right)x\right)$, while that for a suplier is $\kappa(1-\alpha)\left(\left(1+\tau_1^H\right)y(x) - \left(\tau_x^H+\tau_x^F\right)x\right)$. The resulting payoffs for these two agents are $\alpha\left(1+\tau_1^H\right)y(x) - \alpha\left(\tau_x^H+\tau_x^F\right)x$ and $(1-\alpha)(1+\tau_1^H)y(x) - (1-\alpha)(\tau_x^H+\tau_x^F)x$, respectively, just as in the case with customized inputs.

[29] Throughout the extension section, we simply assume that second-order conditions are met.

Nash equilibrium policy choices of governments are affected by conditions of ex-post *bargaining* between foreign suppliers and domestic producers. For instance, the extent to which the foreign country wants to use export taxes to extract revenue from Home producers is tightly related to their bargaining power; in fact, a prerequisite for export taxes to be used by the foreign country is for this country's suppliers to have weak (or at least less-than-complete) bargaining power in the market for inputs.

Leaving aside the differences that the specifics of the bargaining process make to the equilibrium Nash policy choices, we emphasize the key general results that continue to hold regardless of the value of α. First, the international hold-up problem persists and the volume of international input trade is inefficiently low as a consequence. To see this, one can manipulate the first-order conditions and use the expression for $\partial\hat{x}/\partial\tau_x^H$ and $\partial\hat{x}/\partial\tau_x^F$ implied by (47) to derive

$$y'(\hat{x}) = 1 - \frac{\hat{x}}{\partial\hat{x}/\partial\tau_x^H} > 1,$$

which implies that $\hat{x} < x^E$. Second, it is evident that $\hat{\tau}_1^{HN} < 0$ for $\alpha < 1$, and so our model continues to predict that there are distortions in the final good market (p_1^H is too low) that arise as a result of the home-country's attempts to extract bargaining surplus from foreign suppliers. Furthermore, in our discussion of the role and design of trade agreements in section 4, we only used the following properties of the equilibrium: (i) $\bar{W}_{p_x^*}^H + \bar{W}_{p_x^*}^F = 0$, (ii) $\partial p_x^*/\partial\tau_x^H = \partial p_x^*/\partial\tau_x^F - 1$, and (iii) $\partial\hat{x}/\partial\tau_x^H = \partial\hat{x}/\partial\tau_x^F$. It is straightforward to verify that each of these properties continue to hold in this variant of the model with more general bargaining features, and hence our conclusions from that section, as summarized in Propositions 5 and 6 , continue to apply.

It is also possible to use the extended model developed in this section to make a broader point. Up to now we have not taken a stance as to whether the home producer and foreign supplier are vertically related or not. According to the transaction-cost approach to the boundaries of the firm (c.f., Coase, 1936, Williamson, 1985), vertical integration would arise precisely when the hold-up inefficiencies that we have modelled above become large relative to the larger "governance" costs of running an integrated organization. According to that view, our novel rationale for trade agreements would disappear because production of the final good could then be characterized by neoclassical production theory.[30] Nevertheless, the property-rights approach to the theory of the firm (c.f., Grossman and Hart, 1986, Hart and Moore, 1990) has persuasively argued that firm boundaries are better understood as determining the relative bargaining power of producers (via the allocation of residual rights of control inherent in the ownership of productive physical assets) rather than as affecting the space of contracts available to economic agents. Under this interpretation, the rationale for trade agreements that we propose in this paper would very much apply to vertically integrated cross-border production relationships. A crude way to capture the essence of the property-rights theory of the firm in terms of the extended model developed in this

[30] This assumes that home producers could hire foreign suppliers in a competitive market at a given price, which is consistent with the transaction-cost assumption of a frictionless integrated structure (see Grossman and Helpman, 2002, for a general equilibrium treatment).

section would be to associate international outsourcing relationships with a low value of α (the bargaining power of home producers) as compared to the value of α applying to international insourcing relationships. With this interpretation, our finding that the fundamental purpose of a trade agreement does not depend on the value of α then suggests as well that the presence or absence of vertical integration would not alter the fundamental purpose of a trade agreement.[31]

6.2 Ex-Ante Lump-Sum Transfers

Our Benchmark Model rules out ex-ante lump-sum transfers between home producers and foreign suppliers. Although this seems a plausible assumption in our international framework where the promises associated with these transfers may be hard to enforce, it is useful to study the robustness of our results to this assumption. For that purpose, we consider the following modification of *stage* 1 of our Benchmark Model:

stage 1. The unit measure of producers in H and suppliers in F are randomly matched, producing a unit measure of matches. Each producer in H and its matched supplier in F bargain over whether to continue their relationship or not and lump-sum transfers are allowed in the bargaining. This *stage*-1 bargaining is captured by the generalized Nash bargaining solution with weights β and $(1-\beta)$ for the home producer and foreign supplier, respectively, where $\beta \in (0,1)$. If the relationship is terminated, both firms exit; if an agreement is reached, the producer retains the supplier and provides it with a list of customized input specifications.

For simplicity, we assume that the remaining stages of the game are as in the Benchmark Model (and, in particular, all *ex-post* bargains are governed by symmetric Nash bargaining). This implies that at *stage* 1, the home producer and the foreign supplier anticipate that if they reach an agreement, they stand to obtain a joint payoff of

$$\pi^H + \pi^F = \left(1+\tau_1^H\right) y\left(\hat{x}\right) - \left(\tau_x^H + \tau_x^F\right)\hat{x} - \hat{x},$$

where $\hat{x}$ is given by

$$\frac{1}{2}(1+\tau_1^H)y'\left(\hat{x}\right) = 1 + \frac{1}{2}\left(\tau_x^H + \tau_x^F\right).$$

Instead, if an agreement is not reached, both firms exit and are left with a payoff equal to 0. It is straightforward to show that $\pi^H + \pi^F > 0$, which implies that all pairs reach an agreement at *stage* 1. Note, however, that because of the lump-sum transfers, the division of profits between home producers and foreign suppliers is now detached from the ex-post bargaining solution.[32] In

[31] This is not to say that the presence or absence of vertically integrated home producers and foreign suppliers would be irrelevant for the nature of trade agreements. On the contrary, to the extent that international factor ownership associated with vertically integrated multinational firms alters the objective functions of each government, the nature of trade agreements could be very much affected (see Blanchard, 2006). Rather, our point is simply that vertical integration does not by itself obviate the need for a trade agreement to address the international hold-up problem.

[32] Still, the equilibrium level of $\hat{x}$ will be identical to that in the Benchmark Model, since foreign suppliers choose $\hat{x}$ to maximize ex-post payoffs (thus ignoring ex-ante payments).

particular, we have:

$$\begin{aligned}
\pi^H &= \beta\left[\left(1+\tau_1^H\right)y\left(\hat{x}\right)-(\tau_x^H+\tau_x^F)\hat{x}-\hat{x}\right], \text{ and}\\
\pi^F &= (1-\beta)\left[\left(1+\tau_1^H\right)y\left(\hat{x}\right)-(\tau_x^H+\tau_x^F)\hat{x}-\hat{x}\right].
\end{aligned}$$

The values of home and foreign welfare are still given by the same equations as in the Benchmark Model but with these new profit levels π^H and π^F applying.

We can next turn to study the Nash equilibrium policy choices of this variant of the model with lump-sum transfers. Manipulating the first-order conditions related to the choices of $\hat{\tau}_1^{HN}$, $\hat{\tau}_x^{HN}$, and $\hat{\tau}_x^{FN}$ delivers:

$$\begin{aligned}
\hat{\tau}_1^{HN} &= -\frac{(1-\beta)\,\hat{x}\left[\frac{y(\hat{x})}{x}-y'\left(\hat{x}\right)\right]}{\left|\partial D_1/\partial p_1^H\right|},\\
\hat{\tau}_x^{HN} &= -\left(1-\beta\right)\frac{\hat{x}}{\partial\hat{x}/\partial\tau_x^H}-\beta+\hat{\tau}_1^{HN}y'\left(\hat{x}\right), \text{ and}\\
\hat{\tau}_x^{FN} &= -\beta\frac{\hat{x}}{\partial\hat{x}/\partial\tau_x^F}-\left(1-\beta\right).
\end{aligned}$$

Again we emphasize a few general results that continue to hold for any $\beta\in(0,1)$. First, the international hold-up problem persists and the volume of international input trade is inefficiently low as a consequence. To see this, one can manipulate the first-order conditions to derive

$$y'\left(\hat{x}\right)=1-\frac{\hat{x}}{\partial\hat{x}/\partial\tau_x^F}>1,$$

which again implies that $\hat{x}<x^E$. Second, it is evident that $\hat{\tau}_1^{HN}<0$ for $\beta<1$, and so our model continues to predict that there are distortions in the final good market (p_1^H is too low) that arise as a result of the home-country's attempts to extract bargaining surplus from foreign suppliers. Finally, because this extension again satisfies (i) $\bar{W}_{p_x^*}^H+\bar{W}_{p_x^*}^F=0$, (ii) $\partial p_x^*/\partial\tau_x^H=\partial p_x^*/\partial\tau_x^F-1$, and (iii) $\partial\hat{x}/\partial\tau_x^H=\partial\hat{x}/\partial\tau_x^F$, we can again conclude that the conclusions in section 4 on the role and design of trade agreements continue to apply here.

6.3 Other Extensions

We finally outline a few other extensions of our Benchmark Model that we have studied. Due to space constraints we do not report most mathematical details, although they can be found in Antràs and Staiger (2008).

I. Multiple Foreign Countries and Search Costs

In our Benchmark Model, we have restricted our analysis to situations in which home producers can only search for suppliers in F. It is straightforward to show that at least some of our results could be overturned when this restriction is relaxed. To see this, consider the case in which there is a second "foreign" country, denoted by S for "South," with an additional unit measure of potential suppliers identical to those in F. Assume that F and S are identical in every other respect, including preferences, technology and bargaining strength. Under these circumstances and as long as $\tau_x^F > \tau_x^S$, all home producers will prefer to match with southern suppliers over suppliers in F. As a result, the government in F will have an incentive to reduce its export tax below the southern one. Pushing this argument further, it is straightforward to show then that the optimal foreign and southern export taxes that emerge from this variant of the model are negative (i.e., they are subsidies) and Home ends up capturing all the welfare gains from offshoring.[33] As a result, the mix of policies $\hat{\tau}_1^{HN}$ and $\hat{\tau}_x^{HN}$ will be efficient and the rationale for a trade agreement will have vanished.

This example, however, is special in a number of ways. To begin with, the assumption that F and S are symmetric is not innocuous: if one of the two foreign countries has a comparative advantage in supplying inputs, it can (and will) maintain a positive export tax (analogous to "limit pricing" in the case of Bertrand competition among firms), and the result of our Benchmark Model is then preserved. More importantly, the structure of the example above imposes that home producers find a match with probability one, no matter where they search for suppliers. As emphasized by Grossman and Helpman (2005), an important feature of offshoring relationships is the costly search for suitable partners. The same characteristics that make offshoring relationships contractually difficult (i.e., customization, international enforceability of contracts, etc.) preclude the existence of a frictionless competitive market for inputs or for suppliers. In Antràs and Staiger (2008), we explicitly introduce these search frictions and confirm that the central findings of our Benchmark Model are robust to the introduction of multiple foreign countries where inputs may be sourced. We refer the reader to that working paper version for the details.

II. Ad Valorem Tariffs

We have assumed throughout that tariffs on final goods and intermediate inputs are specific. It is straightforward to verify that nothing substantive changes if the final good tariffs are expressed in ad valorem terms. The case of ad valorem import tariff on intermediate inputs is more interesting. In particular, in Antràs and Staiger (2008), we show that ad valorem input tariffs introduce a novel channel through which bargaining between the home producer and foreign supplier can be affected. Despite this novel channel, however, we confirm that the role played by an international trade agreement remains the same.

The key new feature associated with ad-valorem tariffs is that these instruments affect the *slope* of the bargaining frontier between the home producer and the foreign supplier. A positive ad valorem import tariff or export tax makes this *slope* steeper, in the sense that they penalize the producer and supplier for shifting surplus toward the foreign supplier (through a high price p_x^*). On the other hand, a negative ad valorem tariff (an import or export subsidy) makes the slope of

[33] The logic is analogous to that behind the fact that Bertrand competition implies marginal-cost pricing.

the bargaining frontier flatter, thus encouraging transfers of surplus toward the foreign supplier.[34] This constitutes a novel channel through which ad valorem trade taxes can affect the severity of the international hold-up problem. This channel is not present when a specific tariff is instead utilized, because the slope of the bargaining frontier between producer and supplier is -1 independent of the level of the specific tariffs τ_x^H and τ_x^F.

When solving for the Nash equilibrium taxes in this setup, however, we confirm the key inefficiencies existing in our Benchmark Model (see Antràs and Staiger, 2008, for details). Hence, while the mechanisms through which specific and ad valorem tariffs on traded inputs influence the international hold-up problem are distinct, the broad conclusions are similar to those obtained above. Combining this with our earlier observation that the form of the final-good tariff is immaterial, we may conclude that the central findings of our Benchmark Model are robust to the form (ad valorem or specific) that tariffs take, despite the different mechanisms that operate in the two environments.[35]

7 Conclusion

In this paper, we have initiated the study of trade agreements in the presence of offshoring. Our findings indicate that the rise of offshoring is likely to complicate the task of trade agreements for two specific reasons.

First, in the presence of offshoring the mechanism by which countries can shift the costs of intervention on to their trading partners is more complicated and extends to a wider set of policies than is the case when offshoring of customized inputs is not present. And second, the underlying problem that a trade agreement must address in the presence of offshoring varies with the political preferences of member governments, a complication that does not arise in the absence of offshoring.

As a consequence of the first complication, the increasing prevalence of offshoring makes it likely that effective trade agreements must extend their focus beyond the traditional market access concerns: as we have observed, this suggests the need for negotiations that might potentially cover a wide array of internal/domestic measures that have not typically been included in the traditional focus of trade agreements. As a consequence of the second complication, the increasing prevalence of offshoring is likely to make it increasingly difficult for governments to rely on simple

[34] We abstract here from the possibility that firms might engage in transfer-pricing-type behavior in order to avoid trade taxes or collect trade subsidies. In our setting, this amounts to assuming that firms do not have other (non-price) means to transfer surplus between them in their bilateral bargain. If they did have such means, then the price they negotiate would be determined completely by the sign of the trade taxes subject only to the ability of governments to regulate such behavior. Even without such means, the firms in our model do respond to government trade taxes by negotiating different prices, but at least when these firms are taken to be engaged in arms-length transactions this would not be interpreted as transfer pricing in the traditional sense.

[35] It is interesting to observe that the novel channel through which ad valorem tariffs alter the bargaining outcome between home producer and foreign supplier – namely, the slope of the bargaining frontier – also suggests that these policy instruments may have a broader class of applicability with regard to their ability to mitigate international hold-up problems than is the case for specific tariffs. For example, if x were reinterpreted as the unverifiable quality of a fixed unit to be traded, so that tariff policy could not then be conditioned on x, a specific tariff on trade in x would lose its ability to affect the hold-up problem, but an ad valorem tariff would continue to be useful in this regard.

and general rules to help them solve their trade-related problems: this suggests that rules such as reciprocity and non-discrimination, which have been pillars of the multilateral trading system under the GATT/WTO, may become less effective as the prevalence of offshoring rises.

Our paper raises many new questions, both theoretical and empirical. Are international prices best thought of as determined through countless bilateral bargains between buyers and sellers, or rather through anonymous market clearing mechanisms? To the extent that it is the former, Do the trade policy stances of governments in practice have systematic impacts on bargaining outcomes and, through this channel, on trade volumes? Which aspects of the evolving architecture of the GATT/WTO might be best understood from the perspective of the theory we develop here as responses to the rise of offshoring in the world economy? And more generally, does the changing nature of international trade indicate the need for fundamental changes in the nature of regional and multilateral institutions that govern the world trading system? These and related questions strike us as especially fertile areas for further research.

Appendix A

A.1. Second-Order Conditions

In this Appendix we provide a discussion of the second-order conditions of the main tariff setting games developed in the main text.

Second-Best Policy Choices in the Benchmark Model

It is easily verified that the second order conditions associated with the first-order conditions in (8) are satisfied. Simply note that evaluated at the equilibrium, we have

$$
\begin{aligned}
\frac{\partial^2 W^W}{\partial\left(\tau_1^H\right)^2} &= \frac{\partial D_1}{\partial p_1^H} + y''(\hat{x})\left(\frac{\partial \hat{x}}{\partial \tau_1^H}\right)^2 < 0 \\
\frac{\partial^2 W^W}{\partial\left(\tau_x\right)^2} &= y''(\hat{x})\left(\frac{\partial \hat{x}}{\partial \tau_x}\right)^2 < 0 \\
\frac{\partial^2 W^W}{\partial \tau_x \partial \tau_1^H} &= y''(\hat{x})\frac{\partial \hat{x}}{\partial \tau_1^H}\frac{\partial \hat{x}}{\partial \tau_x} > 0
\end{aligned}
$$

and thus $\left(\partial^2 W^W/\partial\left(\tau_1^H\right)^2\right)\left(\partial^2 W^W/\partial\left(\tau_x\right)^2\right) - \left(\partial^2 W^W/\partial \tau_x \partial \tau_1^H\right)^2 = \left(\partial D_1/\partial p_1^H\right) y''(\hat{x})\left(\partial \hat{x}/\partial \tau_x\right)^2 > 0.$

Nash Equilibrium Policy Choices in the Benchmark Model

We now consider the second-order conditions of the Nash equilibrium. Using equation (15), that is $\frac{1}{2}\left(1+\hat{\tau}_1^H\right)y'(\hat{x}) = 1 + \frac{1}{2}\hat{\tau}_x^H + \frac{1}{2}\hat{\tau}_x^F$, we can simplify the first-order conditions to obtain:

$$
\begin{aligned}
\frac{\partial W^H}{\partial \tau_1^H} &= 0 = \tau_1^H \frac{\partial D_1}{\partial p_1^H} - \frac{1}{2}y(\hat{x}) + \left[y'(\hat{x}) - 1 - \tau_x^F\right]\frac{\partial \hat{x}}{\partial \tau_1^H}, \\
\frac{\partial W^H}{\partial \tau_x^H} &= 0 = \frac{1}{2}\hat{x} + \left[y'(\hat{x}) - 1 - \tau_x^F\right]\frac{\partial \hat{x}}{\partial \tau_x^H}, \text{ and} \\
\frac{\partial W^F}{\partial \tau_x^F} &= 0 = \frac{1}{2}\hat{x} + \tau_x^F \frac{\partial \hat{x}}{\partial \tau_x^F}.
\end{aligned}
$$

Consider first the second-order condition for the choice of τ_x^F, i.e., $\partial^2 W^H/\partial\left(\tau_x^F\right)^2 < 0$. Differentiating the last expression above with respect to τ_x^F, we have

$$
\frac{\partial^2 W^F}{\partial\left(\tau_x^F\right)^2} = \frac{3}{2}\frac{\partial \hat{x}}{\partial \tau_x^F} + \tau_x^F \frac{\partial^2 \hat{x}}{\partial\left(\tau_x^F\right)^2}. \tag{A1}
$$

But using the implicit function theorem on (15), we have

$$
\frac{\partial \hat{x}}{\partial \tau_x^F} = \frac{1}{\left(1+\tau_1^H\right)y''(\hat{x})}, \tag{A2}
$$

which implies

$$
\frac{\partial^2 \hat{x}}{\partial\left(\tau_x^F\right)^2} = -\frac{1}{\left(1+\tau_1^H\right)\left(y''(\hat{x})\right)^2}y'''(\hat{x})\frac{\partial \hat{x}}{\partial \tau_x^F}. \tag{A3}
$$

Using these expressions as well as $\tau_x^F = -\frac{1}{2}\hat{x}/\left(\partial\hat{x}/\partial\tau_x^F\right)$, we can write (A1) as

$$\frac{\partial^2 W^F}{\partial\left(\tau_x^F\right)^2} = \frac{1}{2}\frac{\partial\hat{x}}{\partial\tau_x^F}\left(3 + \frac{\hat{x}y'''\left(\hat{x}\right)}{y''\left(\hat{x}\right)}\right),$$

which is negative only if $3+\hat{x}y'''\left(\hat{x}\right)/y''\left(\hat{x}\right) > 0$. As an example, assume that $y\left(x\right) = x^{\eta}/\eta$, with $\eta \in (0,1)$. In such case, we have $y''\left(x\right) = \left(\eta-1\right)x^{\eta-2}$ and $y'''\left(x\right) = \left(\eta-2\right)\left(\eta-1\right)x^{\eta-3}$, and hence $3+\hat{x}y'''\left(\hat{x}\right)/y''\left(\hat{x}\right) = 1+\eta$, which is indeed positive.

The fact that in the Nash equilibrium we have $\hat{\tau}_1^H \neq 0$ implies that the second-order conditions for the choice of τ_1^H and τ_x^H are quite cumbersome to characterize, as they will now also involve properties of the demand function. Throughout the paper, we simply assume that they are satisfied without providing the exact conditions needed.

We next, however, develop a particular case of our model where the second order conditions are easy to characterize and simple comparative statics can be obtained. In particular, we make the simplifying assumption that demand for the final-good is perfectly elastic, which implies that $\hat{\tau}_1^H = 0$. Under this assumption note that it is sufficient to check that $\partial^2 W^H/\partial\left(\tau_x^H\right)^2 < 0$, which requires that

$$\frac{\partial^2 W^H}{\partial\left(\tau_x^H\right)^2} = \frac{1}{2}\frac{\partial\hat{x}}{\partial\tau_x^H} + y''\left(\hat{x}\right)\left(\frac{\partial\hat{x}}{\partial\tau_x^H}\right)^2 + \left[y'\left(\hat{x}\right) - 1 - \tau_x^F\right]\frac{\partial^2\hat{x}}{\partial\left(\tau_x^H\right)^2} < 0.$$

Imposing $\partial W^H/\partial\tau_x^H = 0$ to eliminate τ_x^F and plugging equations (A2) and (A3) – which also apply for τ_x^H –, we can simplify the above expression to:

$$\frac{\partial^2 W^H}{\partial\left(\tau_x^H\right)^2} = \frac{1}{2}\frac{\partial\hat{x}}{\partial\tau_x^H}\left(3 + \frac{\hat{x}y'''\left(\hat{x}\right)}{y''\left(\hat{x}\right)}\right) < 0,$$

This again requires $3+\hat{x}y'''\left(\hat{x}\right)/y''\left(\hat{x}\right) > 0$, which is the same condition as in the choice of τ_x^F.

A.2. Nash Policy Choices with Political Economy

In light of the welfare functions derived in section 5.1, the first-order conditions that define the Nash policies $\hat{\tau}_1^{HN}$, $\hat{\tau}_x^{HN}$, and $\hat{\tau}_x^{FN}$ can be written as (it is straightforward to show that introduction of political economy does not create a reason for F to utilize τ_1^F):

$$\begin{aligned}
\frac{\partial W^H}{\partial\tau_1^H} &= 0 \Rightarrow \tau_1^H\frac{\partial D_1}{\partial p_1^H} + \left(\frac{\gamma^H}{2} - 1\right)\hat{x}\left[\frac{y\left(\hat{x}\right)}{\hat{x}} - y'\left(\hat{x}\right)\right] = 0, \\
\frac{\partial W^H}{\partial\tau_x^H} &= 0 \Rightarrow \left(1 - \frac{\gamma^H}{2}\right)\hat{x} + \left[\gamma^H + \tau_x^H - \tau_1^H y'\left(\hat{x}\right)\right]\frac{\partial\hat{x}}{\partial\tau_x^H} = 0, \text{ and} \\
\frac{\partial W^F}{\partial\tau_x^F} &= 0 \Rightarrow \left(1 - \frac{\gamma^F}{2}\right)\hat{x} + \tau_x^F\frac{\partial\hat{x}}{\partial\tau_x^F} = 0.
\end{aligned} \tag{48}$$

Recalling that, by (15), we have $-\frac{\partial\hat{x}/\partial\tau_1^H}{\partial\hat{x}/\partial\tau_x^H} = y'\left(\hat{x}\right)$, the first condition in (48) can be derived by multiplying the condition $\partial W^H/\partial\tau_x^H = 0$ by $y'\left(\hat{x}\right)$ and adding the resulting expression to the condition $\partial W^H/\partial\tau_1^H = 0$.

Straightforward manipulation of these first-order conditions delivers the following expression for the choice of τ_1^H :

$$\hat{\tau}_1^{HN} = -\frac{\left(1-\frac{\gamma^H}{2}\right)\hat{x}\left[\frac{y(\hat{x})}{\hat{x}} - y'\left(\hat{x}\right)\right]}{\left|\partial D_1/\partial p_1^H\right|}, \tag{49}$$

which naturally reduces to the analogous equation (11) in the Benchmark Model when $\gamma^H = 1$. Notice that for low enough γ^H (in particular $\gamma^H < 2$), the home government continues to find it optimal in the Nash equilibrium to set a positive export tax (or import subsidy) on the final good. Nevertheless, when the weight that the home government places on producer surplus becomes sufficiently high (i.e., $\gamma^H > 2$), $\hat{\tau}_1^{HN}$ flips sign according to (11) and (49) and becomes positive. In such a case, the home government puts in place a Nash trade policy that leads to an *increase* in the domestic price of the final good (i.e., an import tariff or export subsidy). As we have shown above, these policies tend to transfer surplus from the home country to the foreign country, but a sufficiently politically influenced home government is willing to allow this because consumers bear a disproportionate part of the cost of this rent-dissipation.

Further manipulation of the first-order conditions also delivers

$$\hat{\tau}_x^{FN} = \frac{\left(\frac{\gamma^F}{2} - 1\right)\hat{x}}{\partial \hat{x}/\partial \tau_x^F},$$

which indicates that for large enough γ^F (in particular $\gamma^F > 2$), the foreign government no longer sets an export tax in the Nash equilibrium but rather chooses to subsidize exports of intermediate inputs. Intuitively, although a subsidy reduces foreign tariff revenue by an amount which is strictly larger than the amount by which foreign profits increase, a sufficiently politically influenced foreign government weights the latter effect disproportionately more, and thus sets a positive export subsidy in the Nash equilibrium.

As argued in the main text, the fact that the magnitude and even the sign of Nash policies are sensitive to political economy considerations is not particularly surprising: analogous findings are reported for example in Grossman and Helpman (1994) and Bagwell and Staiger (2002, Chapter 10). This suggests that the "positive" predictions of our model regarding the types of instruments that governments will use in equilibrium are not always robust to the presence of political-economy concerns. By the same token, however, the fact that import subsidies and export taxes are rarely observed in the real world should not be interpreted as invalidating the empirical relevance of the trade policy inefficiencies highlighted by our Benchmark Model.

Appendix B: A Competitive Benchmark

For comparison, we now develop the competitive analogue of our (political-economy augmented) model. We suppose that foreign inputs are competitively supplied according to the supply curve

$$x_S^F \equiv x_S^F(p_x^F),$$

In country H, the final good 1 is produced according to the concave production function $y(x)$, and the marginal cost of production of final good 1 is given by

$$mc_1^H = \frac{p_x^H}{y'(x)}.$$

Competitive supply of final good 1 in country H is then determined according to $p_1^H = mc_1^H$ or

$$p_1^H = \frac{p_x^H}{y'(x_D^H)},$$

which implicitly defines x_D^H, the derived demand for the input x, as

$$x_D^H = y'^{-1}\left(p_x^H/p_1^H\right) \equiv x_D^H(p_1^H, p_x^H).$$

The pricing relationships are (with p_x^* the international or world/untaxed price):

$$p_1^H = 1 + \tau_1^H \equiv p_1^H(\tau_1^H); \quad p_x^H = p_x^* + \tau_x^H \equiv p_x^H(\tau_x^H, p_x^*); \quad p_x^F = p_x^* - \tau_x^F \equiv \ p_x^F(\tau_x^F, p_x^*).$$

The market-clearing condition in the world (home and foreign) x market is then given by $x_D^H = x_S^F$, or

$$x_D^H(p_1^H(\tau_1^H), p_x^H(\tau_x^H, p_x^*)) = x_S^F(p_x^F(\tau_x^F, p_x^*)), \tag{B1}$$

which determines $p_x^*(\tau_1^H, \tau_x^H, \tau_x^F)$. Market-clearing input trade volume may then be written as $\hat{x}(p_1^H, p_x^H) \equiv x_D^H(p_1^H(\tau_1^H), p_x^H(\tau_x^H, p_x^*(\tau_1^H, \tau_x^H, \tau_x^F)))$ or equivalently $\hat{x}(p_x^F) \equiv x_S^F(p_x^F(\tau_x^F, p_x^*(\tau_1^H, \tau_x^H, \tau_x^F)))$. We also have $y(p_1^H, p_x^H) \equiv y(\hat{x}(p_1^H, p_x^H))$. Notice that (B1) can be differentiated to yield

$$\frac{\partial p_x^*}{\partial \tau_x^H} = \frac{-\frac{\partial x_D^H(p_1^H, p_x^H)}{\partial p_x^H}}{\frac{\partial x_D^H(p_1^H, p_x^H)}{\partial p_x^H} - \frac{\partial x_S^F(p_x^F)}{\partial p_x^F}} < 0; \quad \frac{\partial p_x^*}{\partial \tau_x^F} = \frac{-\frac{\partial x_S^F(p_x^F)}{\partial p_x^F}}{\frac{\partial x_D^H(p_1^H, p_x^H)}{\partial p_x^H} - \frac{\partial x_S^F(p_x^F)}{\partial p_x^F}} > 0,$$

and so we have that

$$1 = \frac{\partial p_x^*}{\partial \tau_x^F} - \frac{\partial p_x^*}{\partial \tau_x^H}. \tag{B2}$$

The home welfare function may now be written as:

$$W^H = CS(p_1^H) + \gamma^H \int_0^{p_1^H} y(p, p_x^H)dp + (p_1^H - 1)[D_1^H(p_1^H) - y(p_1^H, p_x^H)] + (p_x^H - p_x^*)\hat{x}(p_1^H, p_x^H),$$

or

$$W^H \equiv W^H(p_1^H, p_x^H, p_x^*).$$

Similarly, the foreign welfare function may now be written as:

$$W^F = CS(1) + \gamma^F \int_0^{p_x^F} x_S^F(p)dp + (p_x^* - p_x^F)\hat{x}(p_x^F),$$

or

$$W^F \equiv W^F(p_x^F, p_x^*).$$

Using the fact that $W_{p_x^*}^F = -W_{p_x^*}^H = \hat{x}$, the efficiency frontier is defined by the three conditions:

$$\begin{aligned}
W_{p_x^H}^H + [W_{p_x^H}^H + W_{p_x^F}^F]\frac{\partial p_x^*}{\partial \tau_x^H} &= 0, \\
-W_{p_x^F}^F + [W_{p_x^F}^F + W_{p_x^H}^H]\frac{\partial p_x^*}{\partial \tau_x^F} &= 0, \quad \text{and} \\
W_{p_1^H}^H + [W_{p_x^H}^H + W_{p_x^F}^F]\frac{\partial p_x^*}{\partial \tau_1^H} &= 0.
\end{aligned}$$

Using (B2), it is easy to show that the first two first-order conditions are identical, and therefore determine the sum of τ_x^H and τ_x^F that is consistent with international efficiency.

To further interpret the conditions for efficiency, we multiply the first efficiency condition by $-[\frac{\partial p_x^*/\partial \tau_1^H}{\partial p_x^*/\partial \tau_x^H}]$

and add it to the third efficiency condition, so that we may then restate the two conditions for international efficiency as

$$W^H_{p^H_x} + [W^H_{p^H_x} + W^F_{p^F_x}]\frac{\partial p^*_x}{\partial \tau^H_x} = 0, \text{ and} \quad \text{(B3)}$$
$$W^H_{p^H_1} - W^H_{p^H_x} \cdot \frac{\partial p^*_x / \partial \tau^H_1}{\partial p^*_x / \partial \tau^H_x} = 0.$$

The interpretation of (B3) is as follows. Let us begin with the second efficiency condition. On the left-hand side is the impact on home welfare of (infinitessimal) changes in the *mix* of τ^H_1 and τ^H_x which hold fixed p^*_x – and hence, by (B1) and with τ^F_x and thus $p^F_x(\tau^F_x, p^*_x)$ unchanged, hold fixed as well the level of x^H_D and therefore the equilibrium level of input trade volume $\hat{x}$. Notice, though, that foreign welfare $W^F(p^F_x(\tau^F_x, p^*_x), p^*_x)$ is unaffected by such changes, because p^*_x is held fixed and τ^F_x is not changed and so, as already mentioned, $p^F_x(\tau^F_x, p^*_x)$ is held fixed as well. Hence, the second efficiency condition in (B3) says simply that, at internationally efficient choices of τ^H_1 and τ^H_x, such changes can have no first-order effect on home welfare either. The first efficiency condition in (B3) then ensures that the sum of τ^H_x and τ^F_x achieves the efficient level of p^F_x, and hence the efficient level of input trade volume in light of the mix of τ^H_1 and τ^H_x that the home country employs to deliver the chosen level of p^*_x and (with τ^F_x fixed) p^F_x.

Next consider the Nash policies. The associated first-order conditions are

$$W^H_{p^H_x} + [W^H_{p^H_x} + W^H_{p^*_x}]\frac{\partial p^*_x}{\partial \tau^H_x} = 0, \quad \text{(B4)}$$
$$-W^F_{p^F_x} + [W^F_{p^F_x} + W^F_{p^*_x}]\frac{\partial p^*_x}{\partial \tau^F_x} = 0, \text{ and}$$
$$W^H_{p^H_1} + [W^H_{p^H_x} + W^H_{p^*_x}]\frac{\partial p^*_x}{\partial \tau^H_1} = 0.$$

Using (B2) and $W^F_{p^*_x} = -W^H_{p^*_x}$, the first two Nash first-order conditions can be added together to yield:

$$W^H_{p^H_x} + [W^H_{p^H_x} + W^F_{p^F_x}]\frac{\partial p^*_x}{\partial \tau^H_x} + W^F_{p^*_x} = 0. \quad \text{(B5)}$$

Comparing (B5) to the first efficiency condition in (B3), the difference is the additional term $W^F_{p^*_x} > 0$ on the left-hand side of (B5), which implies that the sum $\tau^H_x + \tau^F_x$ is *inefficiently high* (the first-order condition for efficiency is negative at the Nash taxes), and therefore that the Nash level of input trade volume is inefficiently low in light of the mix of τ^H_1 and τ^H_x that the home country employs in the Nash equilibrium to deliver the chosen level of p^*_x and (with τ^F_x fixed) p^F_x.

Next we multiply the initial first-order condition in (B4) by $-[\frac{\partial p^*_x / \partial \tau^H_1}{\partial p^*_x / \partial \tau^H_x}]$ and add it to the last first-order condition to get

$$W^H_{p^H_1} - W^H_{p^H_x} \cdot \frac{\partial p^*_x / \partial \tau^H_1}{\partial p^*_x / \partial \tau^H_x} = 0. \quad \text{(B6)}$$

Comparing (B6) to the second efficiency condition in (B3), we may conclude that the *mix* of τ^H_1 and τ^H_x that the home country employs in the Nash equilibrium to deliver its chosen level of p^*_x and hence p^F_x – and therefore by (B1), x^H_D and hence $\hat{x}$ – is internationally *efficient* (see Bagwell and Staiger, 2001, for an analogous observation).

Therefore, we may conclude that the single inefficiency in the Nash equilibrium in our competitive benchmark model is that the sum $\tau_x^H + \tau_x^F$ is *inefficiently high*, and hence that there is too little equilibrium input trade volume/input "market access": in the competitive benchmark model, the task of a trade agreement is thus to expand and secure market access to internationally efficient levels (see Bagwell and Staiger, 2001, 2002, for an interpretation of analogous findings from a market access perspective).

Next consider the political optimum conditions. Specifically, consider the hypothetical situation that governments are *not motivated* by the impact of their tariff choices on p_x^*, in the specific sense that $W_{p_x^*}^H \frac{\partial p_x^*}{\partial \tau_1^H} = W_{p_x^*}^H \frac{\partial p_x^*}{\partial \tau_x^H} \equiv 0$ and similarly for W^F. We then identify the tariffs that would be chosen unilaterally (i.e., non-cooperatively) by governments with these hypothetical preferences and ask whether these tariffs are efficient with respect to the actual government preferences. This is Bagwell and Staiger's (1999) original definition, and it is direct to show using (B4) that in our competitive benchmark model the following conditions define the political optimum:

$$W_{p_x^H}^H = 0,\ W_{p_x^F}^F = 0, \text{ and } W_{p_1^H}^H = 0. \tag{B7}$$

Clearly, as an examination of (B3) indicates, the political optimum defined in (B7) is efficient in this setting, whether or not governments are motivated by political economy concerns, so we now have shown that the standard terms-of-trade theory applies in a competitive-supplier version of our set-up.

References

Aghion, Philippe, Pol Antràs, and Elhanan Helpman (2007), "Negotiating Free Trade," *Journal of International Economics*, Vol. 73, No. 1, pp. 1-30.

Antràs, Pol. (2003), "Firms, Contracts, and Trade Structure," *Quarterly Journal of Economics*, 118:4, pp. 1375-1418.

Antràs, Pol (2005), "Incomplete Contracts and the Product Cycle," *American Economic Review*, 95:4, pp. 1054-1073.

Antràs, Pol, Luis Garicano and Esteban Rossi-Hansberg (2006), "Offshoring in a Knowledge Economy," *Quarterly Journal of Economics*, 121:1, pp. 31-77.

Antràs, Pol, and Elhanan Helpman (2004), "Global Sourcing," *Journal of Political Economy* 112:3, pp. 552-580.

Antràs, Pol and Esteban Rossi-Hansberg (2008), "Organizations and Trade," *Annual Review of Economics*, forthcoming.

Bagwell, Kyle and Robert W. Staiger (1999), "An Economic Theory of GATT," *American Economic Review*, Vol. 89, No. 1, pp. 215-248.

Bagwell, Kyle and Robert W. Staiger (2001a), "Domestic Policies, National Sovereignty and International Economic Institutions," *Quarterly Journal of Economics*, Vol. 116, No. 2, pp. 519-562.

Bagwell, Kyle and Robert W. Staiger (2001b), "Reciprocity, non-discrimination and preferential agreements in the multilateral trading system," *European Journal of Political Economy* 17: 281-325.

Bagwell, Kyle and Robert W. Staiger (2002), *The Economics of the World Trading System*, MIT Press.

Bagwell, Kyle and Robert W. Staiger (2006), "What Do Trade Negotiators Negotiate About? Empirical Evidence from the World Trade Organization," NBER Working Paper No. 12727.

Bagwell, Kyle and Robert W. Staiger (2009a), "Delocation and trade agreements in imperfectly competitive markets," Mimeogr., Stanford University.

Bagwell, Kyle and Robert W. Staiger (2009b), "Profit shifting and trade agreements in imperfectly competitive markets," NBER Working Paper Number 14803.

Baron, David (1994), "Electoral Competition with Informed and Uninformed Voters," *American Political Science Review*, 88(1), pp. 33-47.

Bhagwati, Jagdish and V. K. Ramaswami (1963), "Domestic Distortions, Tariffs, and the Theory of Optimum Subsidy," *Journal of Political Economy*, 71:1, pp., 44-50.

Blanchard, Emily (2006), "Foreign Direct Investment, Endogenous Tariffs, and Preferential Trade Agreements," working paper University of Virginia.

Bolton, Patrick and Mathias Dewatripont (2005), *Contract Theory*, MIT Press.

Borga, Maria and William J. Zeile (2004), "International Fragmentation of Production and Intrafirm Trade of U.S. Multinational Companies," Bureau of Economic Analysis Working Paper WP # 2004-02.

Brander, James A. and Barbara J. Spencer (1985), "Export Subsidies and International Market Share Rivalry," *Journal of International Economics*, 18:1-2, pp. 83-100.

Campa, Jose M., and Goldberg, Linda S. (1997), "The Evolving External Orientation of Manufacturing Industries: Evidence from Four Countries." *Federal Reserve Bank of New York Economic Policy Review*, 3 (July), pp. 53-81.

Coase, Ronald H. (1937), "The Nature of the Firm," *Economica*, 4:16, pp. 386-405.

Deardorff, Alan V. (2001), "Fragmentation in Simple Trade Models," *North American Journal of International Economics and Finance*, 12(2), p. 121-137.

Diez, Federico J. (2008), "The Disparate Effects of Tariffs on Offshoring Industries," Mimeo, University of Wisconsin, August.

Economist, The. "The Ins and Outs of Outing." (August 31, 1991): 54-56.

Ethier, Wilfred J. (2004), "Political Externalities, Nondiscrimination, and a Multilateral World," *Review of International Economics*, Vol. 12, No. 3, pp. 303-320.

Feenstra, Robert C. (1998), "Integration of Trade and Disintegration of Production in the Global Economy." *Journal of Economic Perspectives* 12 (Autumn), pp. 31-50.

Feenstra, Robert C., and Gordon H. Hanson (1996a), "Foreign Investment, Outsourcing and Relative Wages," in Robert C. Feenstra, Gene M. Grossman, and Douglas A. Irwin, eds., *Political Economy of Trade Policy: Essays in Honor of Jagdish Bhagwati* (Cambridge. MA: MIT Press), pp. 89-127.

Feenstra, Robert C., and Hanson, Gordon H. (1996b), "Globalization, Outsourcing, and Wage Inequality." *American Economic Review,* 86:2, pp. 240-245.

Feenstra, Robert C., and Hanson, Gordon H. (2005), "Ownership and Control in Outsourcing to China: Estimating the Property Rights Theory of the Firm," *Quarterly Journal of Economics,* 120(2): 729-762.

Grossman, Gene M. and Elhanan Helpman (1994), "Protection for Sale," *American Economic Review*, Vol. 84, No. 4, pp. 833-850.

Grossman, Gene M. and Elhanan Helpman (1996), "Electoral Competition and Special Interest Politics," *Review of Economic Studies*, 63:2, pp. 265-286

Grossman, Gene M. and Elhanan Helpman (2002), "Integration vs. Outsourcing in Industry Equilibrium," *Quarterly Journal of Economics* 117 (1), 85-120.

Grossman, Gene M. and Elhanan Helpman (2005), "Outsourcing in a Global Economy," *Review of Economic Studies*, 72:1, pp. 135-159.

Grossman, Gene M. and Esteban Rossi-Hansberg (2008), "Trading Tasks: A Simple Theory of Offshoring," forthcoming *American Economic Review.*

Grossman, Sanford J., and Oliver D. Hart (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94:4, pp. 691-719.

Hart, Oliver and John Moore (1990), "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98:6, pp. 1119-1158.

Helpman, Elhanan (2006), "Trade, FDI and the Organization of Firms," *Journal of Economic Literature*, 44, pp.589–630.

Hummels, David; Ishii, Jun; and Yi, Kei-Mu. (2001), "The Nature and Growth of Vertical Specialization in World Trade." *Journal of International Economics*, 54, pp. 75-96.

Johnson, Harry G. (1965), "Optimal Intervention in the Presence of Domestic Distortions," in R. E. Caves, P. B. Kenen, and H.G. Johnson, eds., *Trade, Growth and the Balance of Payments: Essays in Honor of Gottfried Haberler*, Amsterdam: North-Holland, pp. 3-34.

Jones, Ronald W. (2000), *Globalization and the Theory of Input Trade*, Cambridge: MIT Press.

Kang, Kichun (2008), "How Much have Export Products Changed from Homogeneous to Differentiated? Evidence from China, Japan, and Korea", *China Economic Review*, Vol. 19 (2), pp. 128-137.

Kemp, Murray C. and Henry Y. Wan, Jr., (1976), "An Elementary Proposition concerning the Formation of Customs Unions," *Journal of International Economics* 6: pp. 95-97.

Levchenko, Andrei (2007), "Institutional Quality and International Trade," *Review of Economic Studies*, 74:3, 791-819.

McLaren, John (1997), "Size, Sunk Costs, and Judge Bowker's Objection to Free Trade," *American Economic Review*, Vol. 87, No. 3, pp. 400-420.

McLaren, John (2000), "Globalization and Vertical Structure," *American Economic Review*, 90, pp. 1239-1254.

Mrazova, Monica (2009), "Trade Negotiations when Market Access Matters," Mimeogr. Oxford University.

Nunn, Nathan (2007), "Relationship-Specificity, Incomplete Contracts and the Pattern of Trade," *Quarterly Journal of Economics*, Vol. 122, No. 2, pp. 569-600.

Nunn, Nathan and Dan Trefler (2008), "The Boundaries of the Multinational Firm: An Empirical Analysis." In *Globalization and the Organization of Firms and Markets*, edited by Elhanan Helpman and Dalia Marin. Cambridge, MA: Harvard University Press, forthcoming.

Ornelas, Emanuel (2008), "Feasible Multilateralism and the Effects of Regionalism," *Journal of International Economics.*, 74:1, pp. 202-224.

Ornelas, Emanuel and John L. Turner (2008a), "Trade Liberalization, Outsourcing, and the Hold Up Problem," *Journal of International Economics*, 74:1, pp. 225-241.

Ornelas, Emanuel and John L. Turner (2008b), "Protection and International Sourcing," Mimeo, July 1.

Ossa, Ralph (2008), "A 'New Trade' Theory of GATT/WTO Negotiations," Mimeo, May.

Ramanarayanan, Ananth (2006), "International Trade Dynamics with Intermediate Inputs," mimeo University of Minnesota.

Rauch, James E. (1999), "Networks versus Markets in International Trade," *Journal of International Economics*, Vol. 48(1), pp. 7-35.

Rosenkranz, Stephanie and Patrick W. Schmitz (2007), "Can Coasean bargaining justify Pigouvian taxation?" *Economica*, vol. 74 (296), pp. 573-585.

Williamson, Oliver E. (1985), *The Economic Institutions of Capitalism*, Free Press: New York.

Yarbrough, Beth V. and Robert M. Yarbrough (1992), *Cooperation and Governance in International Trade: The Strategic Organizational Approach*, Princeton University Press.

Yeaple, Stephen R. (2006), "Offshoring, Foreign Direct Investment, and the Structure of U.S. International Trade," *Journal of the European Economic Association Papers and Proceedings*, 4, pp. 602–611

Yeats, Alexander J. (2001), "Just How Big Is Global Production Sharing?" In Arndt, Sven W., and Kierzkowski, Henry eds., *Fragmentation: New Production Patterns in the World Economy*, Oxford University Press.